PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12Q 1/20

A1

(11) International Publicati n Number: WO 90/14436

(43) International Publication Date: 29 November 1990 (29.11.90)

(21) International Application Number: PCT/US90/02853

(22) International Filing Date: 18 May 1990 (18.05.90)

(30) Priority data:

| | | |
|---|---|---|
| 355,961 | 18 May 1989 (18.05.89) | US |
| 456,637 | 21 December 1989 (21.12.89) | US |
| 505,435 | 4 April 1990 (04.04.90) | US |


(71) Applicant: CHIRON CORPORATION [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US).

(72) Inventors: HOUGHTON, Michael ; 53 Rosemead Court, Danville, CA 94526 (US). CHOO, Qui-Lim ; 5700 Fern Street, El Cerrito, CA 94530 (US). KUO, George ; 1370 Sixth Avenue, San Francisco, CA 94112 (US). HAN, Jang ; 3238 Del Mar, Lafayette, CA 94549 (US). URDEA, Michael, S. ; 100 Bunce Meadow Road, Alamo, CA 94501 (US). WEINER, Amy, J. ; 1766 Indian Way, Oakland, CA 94611 (US).

(74) Agents: MONROY, Gladys, H. et al.; Irell & Manella, 545 Middlefield Road, Suite 200, Menlo Park, CA 94025 (US).


(81) Designated States: AT, AU, BB, BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH, CM (OAPI patent), DE*, DK, ES, FI, GA (OAPI patent), GB, HU, JP, KP, KR, LK, LU, MC, MG, ML (OAPI patent), MR (OAPI patent), MW, NL, NO, RO, SD, SE, SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent).

Published

*With international search report.*

*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*


(54) Title: NANBV DIAGNOSTICS: POLYNUCLEOTIDES USEFUL FOR SCREENING FOR HEPATITIS C VIRUS

(57) Abstract

A new virus, Hepatitis C virus (HCV), which has proven to be the major etiologic agent of blood-borne NANBH, was discovered by Applicant. Reagents for isolating, amplifying, and detecting HCV polynucleotides are provided. These reagents are oligomers comprised of polynucleotide sequences which are capable of forming hybrid structures with HCV target polynucleotide sequences.


* See back of page

## DESIGNATIONS OF "DE"

Until further notice, any designation of "DE" in any international application whose international filing date is prior to October 3, 1990, shall have effect in the territory of the Federal Republic of Germany with the exception of the territory of the former German Democratic Republic.

***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| AT | Austria | ES | Spain | MC | Monaco |
| AU | Australia | FI | Finland | MG | Madagascar |
| BB | Barbados | FR | France | ML | Mali |
| BE | Belgium | GA | Gabon | MR | Mauritania |
| BF | Burkina Fasso | GB | United Kingdom | MW | Malawi |
| BG | Bulgaria | GR | Greece | NL | Netherlands |
| BJ | Benin | HU | Hungary | NO | Norway |
| BR | Brazil | IT | Italy | RO | Romania |
| CA | Canada | JP | Japan | SD | Sudan |
| CF | Central African Republic | KP | Democratic People's Republic of Korea | SE | Sweden |
| CG | Congo | | | SN | Senegal |
| CH | Switzerland | KR | Republic of Korea | SU | Soviet Union |
| CM | Cameroon | LI | Liechtenstein | TD | Chad |
| DE | Germany, Federal Republic of | LK | Sri Lanka | TG | Togo |
| DK | Denmark | LU | Luxembourg | US | United States of America |

# NANBV DIAGNOSTICS: POLYNUCLEOTIDES USEFUL FOR SCREENING FOR HEPATITIS C VIRUS

## Technical Field

The invention relates to materials and methodologies for managing the spread of non-A, non-B hepatitis virus (NANBV) infection. More specifically, it relates to an etiologic agent of non-A, non-B hepatitis (NANBH), hepatitis C virus (HCV), and to polynucleotides and analogs thereof, which are useful in assays for th detection of HCV in biological samples.

## References Cited in the Application

Barr et al. (1986), Biotechniques 4:428.

Beaucage et al. (1981), Tetrahedron Letters 22:1859.

Botstein (1979), Gene 8:17.

Brinton, M.A. (1986) in THE VIRUSES: THE TOGAVIRIDAE AND FLAVIVIRIDAE (Series eds. Fraenkel-Conrat and Wagner, vol. eds. Schlesinger and Schlesinger, Plenum Press), p.327-374.

Broach (1981) in: Molecular Biology of the Yeast Saccharomyces, Vol. 1, p.445, Cold Spring Harbor Press.

Broach et al. (1983), Meth. Enz. 101:307.

Brown et al. (1979), Methods in Enzymology 68:109.

Byrne et al. (1988), Nucleic Acids Res. 16:4165.

Castle et al. (1986), Virology 119:10.

Chang et al. (1977), Nature 198:1056.

Chirgwin et al. (1979), Biochemistry 18:5294.

Choo t 1. (1989), Scienc 244:359.

Chomczynski and Sacchi (1987), Analytical Biochemistry 162:156.

Clewell et al. (1969), Proc. Natl. Acad. Sci. USA 62:1159.

Clewell (1972), J. Bacteriol. 110:667.

Cohen (1972), Proc. Natl. Acad. Sci. USA 69:2110.

Cousens et al. (1987), Gene 61:265.

De Boer et al. (1983), Proc. Natl. Acad. Sci. USA 292:128.

Dreesman et al. (1985), J. Infect. Disease 151:761.

Feinstone et al. (1981), J. Inf. Dis. 144: 588.

Feinstone et al. (1983), Infection and Immunology 41:816.

Feinstone, S.M. and Hoofnagle, J.H. (1984), New Engl. J. Med. 311:185.

Fields & Knipe (1986), FUNDAMENTAL VIROLOGY (Raven Press, N.Y.).

Fiers et al. (1978), Nature 273:113.

Gerety, R.J. et al., in VIRAL HEPATITIS AND LIVER DISEASE (Vyas, B.N., Dienstag, J.L., and Hoofnagle, J.H., eds, Grune and Stratton, Inc., 1984) pp 23-47.

Goeddel et al. (1980), Nucleic Acids Res. 8:4057.

Graham and Van der Eb (1978), Virology 52:546.

Grunstein and Hogness (1975), Proc. Natl. Acad. Sci. USA 73:3961.

Grych et al. (1985), Nature 316:74.

Gubler and Hoffman (1983), Gene 25:263.

Hahn et al. (1988), Virology 162:167.

Han (1987), Biochemistry 26:1617.

Hammerling et al. (1981), MONOCLONAL ANTIBODIES AND T-CELL HYBRIDOMAS.

Hess et al. (1968), J. Adv. Enzyme Reg 7:149.

Hinnen et al. (1978), Proc. Natl. Acad. Sci. 75:1929.

Hitzeman et al. (1980), J. Biol. Chem. 255:2073.

Holland et al. (1978), Biochemistry 17:4900.

Holland (1981), J. Biol. Chem. 256: 1385.

Houghton et al. (1981), Nucleic Acids Res. 9:247

Hunyh, T.V. et al. (1985) in DNA CLONING TECHNIQUES; A

PRACTICAL APPROACH (D. Glov r, Ed., IRL Press, Oxford, U.K.) pp. 49-78.
Immun. R v. (1982) 62:185.
Iwarson (1987), British Medical J. 295:946.
K nnett t al. (1980) MONOCLONAL ANTIBODIES.
Kuo et al. (1989), Science 244:362.
Kyte and Doolittle (1982), J. Mol. Biol. 157:105-132.
Landegren et al. (1988), Science 242:229.
Maniatis, T., et al. (1982) MOLECULAR CLONING; A LABORATORY MANUAL (Cold Spring Harbor Press, Cold Spring Harbor, N.Y.).
Matthews and Kricka (1988), Analytical Biochemistry 169:1.
METHODS IN ENZYMOLOGY (Academic Press).
Mittlin (1989), Clinical chem. 35:1819.
Laemmli (1970), Nature 227, 680.
Lee et al. (1988), Science 239:1288.
Loh et al. (1989), Science 243:217.
Mackow et al. (1987), Virology 159:217.
Mayer and Walker, eds. (1987), IMMUNOCHEMICAL METHODS IN CELL AND MOLECULAR BIOLOGY (Academic Press, London).
Maxam et al. (1980), Methods in Enzymology 65:499.
MacNamara et al. (1984), Science 226:1325.
Messing et al. (1981), Nucleic Acids Res. 9:309.
Messing (1983), Methods in Enzymology 101:20-37.
METHODS IN ENZYMOLOGY (Academic Press).
Michelle et al., Int. Symposium on Viral Hepatitis.
Monath (1986) in THE VIRUSES: THE TOGAVIRADAE AND FLAVIVIRIDAE (Series eds. Fraenkel-Conrat and Wagner, vol. eds. Schlesinger and Schlesinger, Plenum Press), p.375-440.
Murakawa et al. (1988), DNA 7:287.
Nagahuma et al. (1984), Anal. Biochem. 141:74.
Narang et al (1979), Methods in Enzymology 68:90.
Neurath et al. (1984), Science 224:392.
Nisonoff et al. (1981), Clin. Immunol. Immunopathol. 21:397-406.

Overby, L.R. (1985), Curr. Hepatol. 5:49.

Overby, L.R. (1986), Curr. Hepatol. 6:65.

Overby, L.R. (1987), Curr. Hepatol. 7:35.

P leg (1969), Nature 221:193.

Pfefferkorn and Shapiro (1974), in COMPREHENSIVE VIROLOGY, Vol. 2 (Fraenkel-Conrat & Wagner, eds., Plenum, N.Y.) pp. 171-230.

Prince, A.M. (1983), Annu. Rev. Microbiol. 37:217.

Rice et al. (1985), Science 229:726.

Rice et al. (1986) in THE VIRUSES: THE TOGAVIRIDAE AND FLAVIVIRIDAE (Series eds. Fraenkel-Conrat and Wagner, vol. eds. Schlesinger and Schlesinger, Plenum Press), p.279-328.

Roehrig (1986) in THE VIRUSES: THE TOGAVIRIDAE AND FLAVIVIRIDAE (Series eds. Fraenkel-Conrat and Wagner, vol. eds. Schlesinger and Schlesinger, Plenum Press)

Rosenberg et al. (1984), Nature 312:7.

Sadler et al. (1980), Gene 8, 279.

Saiki et al. (1985), Science 230:1350.

Saiki et al. (1986), Nature 324: 163.

Saiki et al. (1988), Science 239:487.

Sanger et al. (1977), Proc. Natl. Acad. Sci. USA 74:5463.

Scharf et al. (1986), Science 233:1076.

Schlesinger et al. (1986), J. Virol. 60:1153.

Schreier, M., et al. (1980) HYBRIDOMA TECHNIQUES

Scopes (1984), PROTEIN PURIFICATION, PRINCIPLES AND PRACTICE, SECOND EDITION (Springer-Verlag, N.Y.).

Shimatake et al. (1981), Nature 292:128.

Shigekawa and Dower (1988), BioTechniques 6:742.

Steimer et al. (1986), J. Virol. 58:9.

Stollar (1980), in THE TOGAVIRUSES (R.W. Schlesinger, ed., Academic Press, N.Y.), pp. 584-622.

Sumiyoshi et al. (1987), Virology 161:497.

Taylor et al. (1976), Biochem. Biophys. Acta 442:324.

Towbin et al. (1979), Proc. Natl. Acad. Sci. USA 76, 4350.

Tsu and Herzenberg (1980), in SELECTED METHODS IN CELLULAR

IMMUNOLOGY (W.H. Freeman and Co.) pp. 373-391.
Vytd haag t al. (1985), J. Immunol. 134:1225.
Val nzu la, P., t al. (1982), Natur 298:344.
Valenzuela, P., et al. (1984), in HEPATITIS B (Millman, I., et al., ed, Plenum Press) pp. 225-236.
Warner (1984), DNA 3:401.
Wu and Grossman (1987), Methods in Enzymology Vol. 154, RECOMBINANT DNA, Part E.
Wu (1987), Methods in Enzymology vol 155, RECOMBINANT DNA, part F.
Zoller (1982), Nucleic Acids Res. 10:6487.

Cited Patents

U.S. Patent No. 4,341,761
U.S. Patent No. 4,399,121
U.S. Patent No. 4,427,783
U.S. Patent No. 4,444,887
U.S. Patent No. 4,466,917
U.S. Patent No. 4,472,500
U.S. Patent No. 4,491,632
U.S. Patent No. 4,493,890
U.S. Patent No. 4,683,202
U.S. Patent No. 4,458,066
U.S. Patent No. 4,868,105

Background Art

Non-A, Non-B hepatitis (NANBH) is a transmissible disease or family of diseases that are believed to be viral-induced, and that are distinguishable from other forms of viral-associated liver diseases, including that caused by the known hepatitis viruses, i.e., hepatitis A virus (HAV), hepatitis B virus (HBV), and delta hepatitis virus (HDV), as well as the hepatitis induced by cytomegalovirus (CMV) or Epstein-Barr virus (EBV). NANBH was first identified in transfused individuals. Transmission from man to chimpanzee and se-

rial passage in chimpanz es provided evidence that NANBH is due to a transmissible infectious agent or agents.

Epidemiologic evid nce is suggestiv that there may be three types of NANBH: the water-borne epidemic type; the blood or needle associated type; and the sporadically occurring (community acquired) type. However, the number of agents which may be the causative of NANBH are unknown.

There have been a number of candidate NANBV. See, for example the reviews by Prince (1983), Feinstone and Hoofnagle (1984), and Overby (1985, 1986, 1987) and the article by Iwarson (1987). However, there is no proof that any of these candidates represent the etiological agent of NANBH.

The demand for sensitive, specific methods for screening and identifying carriers of NANBV and NANBV contaminated blood or blood products is significant. Post-transfusion hepatitis (PTH) occurs in approximately 10% of transfused patients, and NANBH accounts for up to 90% of these cases. The major problem in this disease is the frequent progression to chronic liver damage (25-55%).

Patient care as well as the prevention of transmission of NANBH by blood and blood products or by close personal contact require reliable screening, diagnostic and prognostic tools to detect nucleic acids, antigens and antibodies related to NANBV.

Methods for detecting specific polynucleotides by hybridization assays are known in the art. See, for example, Matthews and Kricka (1988), Analytical Biochemistry 169:1; Landegren et al. (1988), Science 242:229; and Mittlin (1989), Clinical chem. 35:1819. U.S. Patent No. 4,868,105, issued Sept. 9, 1989, and in E.P.O. Publication No. 225,807 (published June 16, 1987).

Applicant discovered a new virus, the Hepatitis C virus (HCV), which has proven to be the major etiologic agent of blood-borne NANBH (BB-NANBH). Applicant's

initial work, including a partial genomic s qu nc of th prototyp HCV isolate, CDC/HCV1 (also called HCV1), is d scribed in E.P.O. Publication No. 318,216 (published 31 May 1989) and PCT Pub. No. WO 89/04669 (publish d 1 June 1989). Th disclosur s of these pat nt applications, as well as any corresponding national patent applications, are incorporated herein by reference. These applications teach, inter alia, recombinant DNA methods of cloning HCV sequences, HCV probe diagnostic techniques, anti-HCV antibodies, and methods of isolating new HCV sequences.

Disclosure of the Invention

The present invention is based on HCV sequences described in E.P.O. Publication No. 318,216 and in PCT Pub. No. WO 89/04669, as well as other HCV sequences that are described herein. Methods for isolating and/or detecting specific polynucleotides by hybridization could not be used for screening for HCV until Applicants' discovery of HCV. Accordingly, one aspect of the invention is an oligomer capable of hybridizing to an HCV sequence in an analyte polynucleotide strand, wherein th oligomer is comprised of an HCV targeting sequence complementary to at least 4 contiguous nucleotides of HCV cDNA shown in Fig. 18.

Another aspect of the invention is a process for detecting an HCV sequence in an analyte strand suspected of containing an HCV polynucleotide, wherein the HCV polynucleotide comprises a selected target region, said process comprising:

(a) providing an oligomer capable of hybridizing to an HCV sequence in an analyte polynucleotide strand, wherein the oligomer is comprised of an HCV targeting sequence complementary to at least 4 contiguous nucleotides of HCV cDNA shown in Fig. 18

(b) incubating the analyte strand with the oligomer of (a) which allow specific hybrid duplexes to

form between the targeting sequence and the target sequence; and

(d) detecting hybrids formed b tween targ t r gion, if any, and the oligomer.

Yet another aspect of th invention is a method for preparing blood free of HCV comprising:

(a) providing analyte nucleic acids from a sample of blood suspected of containing an HCV target sequence;

(b) providing an oligomer capable of hybridizing to the HCV sequence in an analyte polynucleotide strand, if any, wherein the oligomer is comprised of an HCV targeting sequence complementary to a sequence of at least 8 nucleotides present in a conserved HCV nucleotide sequence in HCV RNA;

(c) reacting (a) with (b) under conditions which allow the formation of a polynucleotide duplex between the targeting sequence and the target sequence, if any;

(d) detecting a duplex formed in (c), if any; and

(e) saving the blood from which complexes wer not detected in (d).

Brief Description of the Drawings

Fig. 1 shows the sequence of the HCV cDNA in clone 12f, and the amino acids encoded therein.

Fig. 2 shows the HCV cDNA sequence in clone k9-1, and the amino acids encoded therein.

Fig. 3 shows the sequence of clone 15e, and the amino acids encoded therein.

Fig. 4 shows the nucleotide sequence of HCV cDNA in clone 13i, the amino acids encoded therein, and the sequences which overlap with clone 12f.

Fig. 5 shows th nucleotide s quence of HCV cDNA in clone 26j, the amino acids encoded therein, and th s qu nces which overlap clone 13i.

Fig. 6 shows the nucleotide sequence of HCV cDNA in clone CA59a, th amino acids encoded therein, and the sequences which overlap with clones 26j and K9-1.

Fig. 7 shows the nucleotide sequence of HCV cDNA in clone CA84a, the amino acids encoded therein, and the sequences which overlap with clone CA59a.

Fig. 8 shows the nucleotide sequence of HCV cDNA in clone CA156e, the amino acids encoded therein, and the sequences which overlap with CA84a.

Fig. 9 shows the nucleotide sequence of HCV cDNA in clone CA167b, the amino acids encoded therein, and the sequences which overlap CA156e.

Fig. 10 shows the nucleotide sequence of HCV cDNA in clone CA216a, the amino acids encoded therein, and the overlap with clone CA167b.

Fig. 11 shows the nucleotide sequence of HCV cDNA in clone CA290a, the amino acids encoded therein, and the overlap with clone CA216a.

Fig. 12 shows the nucleotide sequence of HCV cDNA in clone ag30a and the overlap with clone CA290a.

Fig. 13 shows the nucleotide sequence of HCV cDNA in clone CA205a, and the overlap with the HCV cDNA sequence in clone CA290a.

Fig. 14 shows the nucleotide sequence of HCV cDNA in clone 18g, and the overlap with the HCV cDNA sequence in clone ag30a.

Fig. 15 shows the nucleotide sequence of HCV cDNA in clone 16jh, the amino acids encoded therein, and the overlap of nucleotides with the HCV cDNA sequence in clone 15e.

Fig. 16 shows the nucleotide sequence of HCV cDNA in clone 6k, the amino acids encoded therein, and the

overlap of nucleotides with the HCV cDNA sequence in clone 16jh.

Fig. 17 shows th nucleotide sequence of HCV cDNA in clone p131jh, the amino acids encoded therein, and th overlap of nucleotides with the HCV cDNA sequence in clone 6k.

Fig. 18 shows the the compiled HCV cDNA sequenc derived from the clones described herein and from the compiled HCV cDNA sequence presented in E.P.O. Publication No. 318,216. The clones from which the sequence was derived are 5'-clone32, b114a, 18g, ag30a, CA205a, CA290a, CA216a, pi14a, CA167b, CA156e, CA84a, CA59a, K9-1 (also called k9-1),26j, 13i, 12f, 14i, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g, 39c, 35f, 19g, 26g, 15e, b5a, 16jh, 6k, and p131jh. In the figure the three horizontal dashes above the sequenc indicate the position of the putative initiator methionine codon. Also shown in the figure is the amino acid sequence of the putative polyprotein encoded in the HCV cDNA. Heterogeneities in cloned DNAs of HCV1 are indicated by the amino acids indicated above the putatively encoded sequence of the large ORF; the parentheses indicate that the heterogeneity was detected at or near to the 5'- or 3'- end of the HCV cDNA in the clone.

Fig. 19 shows the sequences of capture and label probes for the detection of HCV RNA in biological samples.

Fig. 20 shows schematic alignment of a flaviviral polyprotein and a putative HCV polyprotein encoded in the major ORF of the HCV genome. Also indicated in the figure are the possible functions of the flaviviral polypeptides cleaved from the flaviral polyprotein. In addition, the relative placements of the HCV polypeptides, $NANB_{5-1-1}$ and C100, with respect to the putative HCV polyprotein are indicated.

Fig. 22 shows th double-stranded nucleotide s qu nc of th HCV cDNA insert in clon 81, and th putativ amino acid sequence of th polypeptide encoded th r in.

Fig. 23 shows th HCV cDNA s quence in clone 36, the segment which overlaps the NANBV cDNA of clone 81, and the polypeptide sequence encoded within clone 36.

Fig. 24 shows the HCV cDNA sequence in clone 37b, the segment which overlaps clone 35, and the polypeptide encoded therein.

Fig. 25 shows autoradiographs of the HCV cPCR assay on RNA derived from liver samples of chimpanzees with NANBH (Fig. 25A) and on Italian patients with NANBH (Fig. 25B).

Fig. 26A and 26B are graphs showing the temporal relationship between the display of liver damage, the presence of HCV RNA, and the presence of anti-HCV antibodies for two chimpanzees with NANBH.

Fig. 27 shows the nucleotide sequence of HCV cDNA in clone CA84a, the amino acids encoded therein, and the sequences which overlap with clone CA59a.

Fig. 28 shows the HCV cDNA sequence in clone 40b, the segment which overlaps clone 37b, and the polypeptide encoded therein.

Fig. 29 is an autoradiograph showing the labeled amplified products of approximately 300, 30, and 3 CID of HCV genomes.

Fig. 32 shows the nucleotide sequence of HCV cDNA in clone 40a.

Fig. 33 is an autoradiograph showing amplified products extended from primers derived from conserved regions of the HCV genome.

Fig. 34 shows the HCV cDNA sequence in clone 35, the segment which overlaps clone 36, and the polypeptide encoded therein.

Fig. 37 is a diagram showing the r lationship of probes and primers derived from the 5'-r gion of HCV RNA, from which th HCV cDNAs in clones ag30a and k9-1 are d rived.

Fig. 38 is an autoradiograph of amplified products extended from sets of primers derived from ag30a and k9-1.

Fig. 39 shows the aligned nucleotide sequences of human isolates 23 and 27 and of HCV1. Homologous sequences are indicated by the symbol (*). Non homologous sequences are in small letters.

Fig. 40 shows the aligned amino acid sequences of human isolates 23 and 27 and of HCV1. Homologous sequences are indicated by the symbol (*). Non homologous sequences are in small letters.

Fig. 41 shows a half-tone reproduction of an autoradiograph of a Northern blot of RNA isolated from the liver of a BB-NANBV infected chimpanzee, probed with BB-NANBV cDNA of clone 81.

Fig. 43 shows a half-tone reproduction of an autoradiograph of nucleic acids extracted from NANBV particles captured from infected plasma with anti-$NANB_{5-1-1}$, and probed with $^{32}P$-labeled NANBV cDNA from clone 81.

Fig. 44 shows reproductions of autoradiographs of filters containing isolated NANBV nucleic acids, probed with $^{32}P$-labeled plus and minus strand DNA probes derived from NANBV cDNA in clone 81.

Fig. 46 shows the nucleotide consensus sequence of human isolate 23, variant sequences are shown below the sequence line. The amino acids encoded in the consensus sequence are also shown.

Fig. 47 shows the nucleotide consensus sequence of human isolate 27, variant sequences are shown below the sequence line. The amino acids encoded in the consensus sequence are also shown.

Fig. 48 is a graph showing th relationship of the EnvL and EnvR primers to th mod l flavivirus polyprotein and putative HCV polyprotein.

Fig. 49 shows a comparison of th composit aligned nucleotide sequences of isolates Thorn, EC1, HCT #18, and HCV1.

Fig. 50 shows a comparison of the nucleotide sequences of EC10 and a composite of the HCV1 sequence; the EC10 sequence is on the line above the dots, and the HCV1 sequence is on the line below the dots.

Fig. 51 shows a comparison of the amino acid sequences 117-308 (relative to HCV1) encoded in the "EnvL" regions of the consensus sequences of human isolates HCT #18, JH23, JH 27, Thorne, EC1, and of HCV1.

Fig. 52 shows a comparison of the amino acid sequences 330-360 (relative to HCV1) encoded in the "EnvR" regions of the consensus sequences of human isolates HCT #18, JH23, JH 27, Thorne, EC1, and of HCV1.

Fig. 53 shows the nucleotide sequences of individual primers in primer mixture 5'-3.

Modes for Carrying Out the Invention

The term "hepatitis C virus" (HCV) has been reserved by workers in the field for an heretofore unknown etiologic agent of NANBH. The prototype isolate of HCV has been identified in U.S.S.N. 122,714 (See also E.P.O. Publication No. 318,216). The term HCV also includes new isolates of the same viral species. As an extension of this terminology, the disease caused by HCV, formerly called blood-borne NANB hepatitis (BB-NANBH), is called hepatitis C. The terms NANBH and hepatitis C may be used interchangeably herein.

HCV is a viral species of which pathogenic strains cause BB-NANBH. There may also be attenuated strains or defective interfering particles derived therefrom. As shown infra, the HCV genome is comprised of

RNA. It is known that RNA containing virus s have relatively high rates of spontaneous mutation, i.e., r portedly on the order of $10^{-3}$ to $10^{-4}$ per incorporated nucl otide (Fields & Knipe (1986)). Therefore, since heterogeneity and fluidity of genotype are inherent in RNA viruses, there are multiple strains/isolates, which may be virulent or avirulent, within the HCV species. The compositions and methods described herein, enable the propagation, identification, detection, and isolation of the various HCV strains or isolates.

Several different strains/isolates of HCV have been identified. (See infra). One such strain or isolate, which is a prototype, is named CDC/HCV1 (also called HCV1). Information from one strain or isolate, such as a partial genomic sequence, is sufficient to allow those skilled in the art using standard techniques to isolate new strains/isolates and to identify whether such new strains/isolates are HCV. For example, several different strains/isolates are described infra. These strains, which were obtained from a number of human sera (and from different geographical areas), were isolated utilizing the information from the genomic sequence of HCV1.

Using the techniques described in E.P.O. Publication No. 318,216 and infra, the genomic structure and the nucleotide sequence of HCV1 genomic RNA has been deduced. The genome appears to be single-stranded RNA containing ~10,000 nucleotides. The genome is positive-stranded, and possesses a continuous, translational open reading frame (ORF) that encodes a polyprotein of about 3,000 amino acids. In the ORF, the structural protein(s) appear to be encoded in approximately the first quarter of the N-terminus region, with the majority of the polyprotein responsible for non-structural proteins. When compared with all known viral sequences, small but significant co-linear homologies are observed with the

non-structural proteins of th flavivirus family, and with the pestiviruses (which ar now also considered to be part of th Flavirus family).

A schematic alignment of possible regions of a flaviviral polyprotein (using Yellow Fever Virus as an example), and of a putative polyprotein encoded in the major ORF of the HCV genome, is shown in Fig. 20. In the figure the possible domains of the HCV polyprotein are indicated. The flavivirus polyprotein contains, from the amino terminus to the carboxy terminus, the nucleocapsid protein (C), the matrix protein (M), the envelope protein (E), and the non-structural proteins (NS) 1, 2 (a+b), 3, 4 (a+b), and 5. Based upon the putative amino acids encoded in the nucleotide sequence of HCV1, a small domain at the extreme N-terminus of the HCV polyprotein appears similar both in size and high content of basic residues to the nucleocapsid protein (C) found at the N-terminus of flaviviral polyproteins. The non-structural proteins 2,3,4, and 5 (NS2-5) of HCV and of yellow fever virus (YFV) appear to have counter parts of similar size and hydropathicity, although there is divergence of the amino acid sequences. However, the region of HCV which would correspond to the regions of YFV polyprotein which contains the M, E, and NS1 protein not only differs in sequence, but also appears to be quite different both in size and hydropathicity. Thus, while certain domains of the HCV genome may be referred to herein as, for example, NS1, or NS2, it should be borne in mind that these designations are speculative; there may be considerable differences between the HCV family and flaviviruses that have yet to be appreciated.

Different strains, isolates or subtypes of HCV are expected to contain variations at the amino acid and nucleic acids compared with HCV1. Many isolates are expected to show much (i.e., more than about 40%) homology in the total amino acid sequence compared with HCV1.

However, it may also be found that there are oth r less homologous HCV isolat s. These would be defined as HCV according to various crit ria such as, for example, an ORF of approximately 9,000 nucleotides to approximately 12,000 nucleotides, encoding a polyprotein similar in size to that of HCV1, an encoded polyprotein of similar hydrophobic and/or antigenic character to that of HCV1, and the presence of co-linear peptide sequences that are conserved with HCV1. In addition, it is believed that the genome would be a positive-stranded RNA.

All HCV isolates encode at least one epitope which is immunologically identifiable (i.e., immunologically cross-reactive) with an epitope encoded in the HCV cDNAs described herein. Preferably the epitope is contained in an amino acid sequence described herein and is unique to HCV when compared to previously known pathogens. The uniqueness of the epitope may be determined by its immunological reactivity with anti-HCV antibodies and lack of immunological reactivity with antibodies to known pathogens.

HCV strains and isolates are evolutionarily related. Therefore, it is expected that the overall homology of the genomes at the nucleotide level may be about 40% or greater, probably will be about 50% or greater, probably about 60% or greater, and even more probably about 80% or greater; and in addition that there will be corresponding contiguous sequences of at least about 13 nucleotides. It should be noted, as shown infra, that there are variable and hypervariable regions within the HCV genome; therefore, the homology in these regions is expected to be significantly less than that in the overall genome. The correspondence between the putative HCV strain genomic sequence and, for example, the CDC/HCV1 cDNA sequence can be determined by techniques known in the art. For example, they can be determined by a direct comparison of the sequence information of the

polynucl otide from the putative HCV, and the HCV cDNA sequence(s) described herein. Th y also can b determin d by hybridization of the polynucleotid s under conditions which form stable duplexes between homologous regions (for example, those which would be used prior to $S_1$ digestion), followed by digestion with single stranded specific nuclease(s), followed by size determination of the digested fragments.

Because of the evolutionary relationship of th strains or isolates of HCV, putative HCV strains or isolates are identifiable by their homology at the polypeptide level. Generally, HCV strains or isolates are expected to be at least 40% homologous, more than about 50% homologous, probably more than about 70% homologous, and even more probably more than about 80% homologous, and some may even be more than about 90% homologous at the polypeptide level. The techniques for determining amino acid sequence homology are known in the art. For example, the amino acid sequence may be determined directly and compared to the sequences provided herein. Alternatively the nucleotide sequence of the genomic material of the putative HCV may be determined (usually via a cDNA intermediate), the putative amino acid sequence encoded therein can be determined, and the corresponding regions compared.

As used herein, a polynucleotide "derived from" a designated sequence refers to a polynucleotide sequence which is comprised of a sequence of approximately at least about 6 nucleotides, preferably at least about 8 nucleotides, more preferably at least about 10-12 nucleotides, and even more preferably at least about 15-20 nucleotides corresponding to a region of the designated nucleotide sequence. "Corresponding" means homologous to or complementary to the designated sequence. Preferably, the sequence of the region from which the polynucleotide is derived is homologous to or complementary to a sequence which is unique to an HCV genome. More preferably, the

derived sequ nc is homologous or complementary to a sequence that is unique to all or to a majority of HCV isolat s. Whether or not a s quence is unique to the HCV g nome can be determined by techniques known to those of skill in the art. For example, the sequence can be compared to sequences in databanks, e.g., Genebank, to determine whether it is present in the uninfected host or other organisms. The sequence can also be compared to the known sequences of other viral agents, including those which are known to induce hepatitis, e.g., HAV, HBV, and HDV, and to members of the Flaviviridae. The correspondence or non-correspondence of the derived sequence to other sequences can also be determined by hybridization under the appropriate stringency conditions. Hybridization techniques for determining the complementarity of nucleic acid sequences are known in the art, and are discussed infra. See also, for example, Maniatis et al. (1982). In addition, mismatches of duplex polynucleotid s formed by hybridization can be determined by known techniques, including for example, digestion with a nuclease such as S1 that specifically digests single-stranded areas in duplex polynucleotides. Regions from which typical DNA sequences may be "derived" include but are not limited to, for example, regions encoding specific epitopes, as well as non-transcribed and/or non-translated regions.

The derived polynucleotide is not necessarily physically derived from the nucleotide sequence shown, but may be generated in any manner, including for example, chemical synthesis or DNA replication or reverse transcription or transcription. In addition, combinations of regions corresponding to that of the designated sequence may be modified in ways known in the art to be consistent with an intended use.

The term "recombinant polynucleotide" as used herein intends a polynucleotide of genomic, cDNA,

semisynth tic, or synth tic origin which, by virtue of its origin or manipulation: (1) is not associated with all or a portion f a polynucl otide with which it is associated in nature, (2) is linked to a polynucleotide other than that to which it is linked in nature, or (3) does not occur in nature.

The term "polynucleotide" as used herein refers to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. This term refers only to the primary structure of the molecule. Thus, this term includes double- and single-stranded DNA and RNA. It also includes known types of modifications, for example, labels which are known in the art, methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as, for example, those with uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoamidates, carbamates, etc.) and with charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), those containing pendant moieties, such as, for example proteins (including for e.g., nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), those with intercalators (e.g., acridine, psoralen, etc.), those containing chelators (e.g., metals, radioactive metals, boron, oxidative metals, etc.), those containing alkylators, those with modified linkages (e.g., alpha anomeric nucleic acids, etc.), as well as unmodified forms of the polynucleotide.

As used herein, the "sense strand" of a nucleic acid contains the sequence that has sequence homology to that of mRNA. The "anti-sense strand" contains a sequence which is complementary to that of the "sense strand".

As used herein, a "positive stranded genome" of a virus is one in which the genome, whether RNA or DNA, is single-stranded and which encodes a viral polypeptide(s). Examples of positive stranded RNA viruses include

Togavirida , Cor navirida , R troviridae, Picornaviridae, and Calicivirida . Included also, are the Flaviviridae, which were formerly classified as Togaviradae. Se Fields & Knipe (1986).

The term "primer" as used herein refers to an oligomer which is capable of acting as a point of initiation of synthesis of a polynucleotide strand when placed under appropriate conditions. The primer will be completely or substantially complementary to a region of the polynucleotide strand to be copied. Thus, under conditions conducive to hybridization, the primer will anneal to the complementary region of the analyte strand. Upon addition of suitable reactants, (e.g., a polymerase, nucleotide triphosphates, and the like), the primer is extended by the polymerizing agent to form a copy of the analyte strand. The primer may be single-stranded, or alternatively may be partially or fully double-stranded.

The terms "analyte polynucleotide" and "analyte strand" refer to a single- or double-stranded nucleic acid molecule which is suspected of containing a target sequence, and which may be present in a biological sample.

As used herein, the term "oligomer" refers to primers and to probes. The term oligomer does not connote the size of the molecule. However, typically oligomers are no greater than 1000 nucleotides, more typically are no greater than 500 nucleotides, even more typically are no greater than 250 nucleotides; they may be no greater than 100 nucleotides, and may be no greater than 75 nucleotides, and also may be no greater than 50 nucleotides in length.

As used herein, the term "probe" refers to a structure comprised of a polynucleotide which forms a hybrid structure with a target sequence, due to complementarity of at least one sequence in the probe with a sequence in the target region. The polynucleotide regions of probes may be composed of DNA, and/or RNA, and/

or synthetic nucleotide analogs. Included within probes ar "captur probes" and "label probes". Pr ferably the probe do s not contain a sequence complementary to sequenc (s) used to prime the polymerase chain reaction (PCR).

As used herein, the term "target region" refers to a region of the nucleic acid which is to be amplified and/or detected. The term "target sequence" refers to a sequence with which a probe or primer will form a stable hybrid under desired conditions.

The term "capture probe" as used herein refers to a polynucleotide comprised of a single-stranded polynucleotide coupled to a binding partner. The single-stranded polynucleotide is comprised of a targeting polynucleotide sequence, which is complementary to a target sequence in a target region to be detected in the analyte polynucleotide. This complementary region is of sufficient length and complementarity to the target sequence to afford a duplex of stability which is sufficient to immobilize the analyte polynucleotide to a solid surface (via the binding partners). The binding partner is specific for a second binding partner; the second binding partner can be bound to the surface of a solid support, or may be linked indirectly via other structures or binding partners to a solid support.

The term "targeting polynucleotide sequence" as used herein, refers to a polynucleotide sequence which is comprised of nucleotides which are complementary to a target nucleotide sequence; the sequence is of sufficient length and complementarity with the target sequence to form a duplex which has sufficient stability for the purpose intended.

The term "binding partner" as used herein refers to a molecule capable of binding a ligand molecule with high specificity, as for example an antigen and an antibody specific therefor. In general, the specific binding

partners must bind with sufficient affinity to immobiliz the analyte copy/complementary strand duplex (in the case of capture probes) under the isolation conditions. Specific binding partners are known in th art, and include, for example, biotin and avidin or streptavidin, IgG and protein A, the numerous known receptor-ligand couples, and complementary polynucleotide strands. In th case of complementary polynucleotide binding partners, th partners are normally at least about 15 bases in length, and may be at least 40 bases in length; in addition, they have a content of Gs and Cs of at least about 40% and as much as about 60%. The polynucleotides may be composed of DNA, RNA, or synthetic nucleotide analogs.

The term "coupled" as used herein refers to attachment by covalent bonds or by strong non-covalent interactions (e.g., hydrophobic interactions, hydrogen bonds, etc.). Covalent bonds may be, for example, ester, ether, phosphoester, amide, peptide, imide, carbon-sulfur bonds, carbon-phosphorus bonds, and the like.

The term "support" refers to any solid or semi-solid surface to which a desired binding partner may be anchored. Suitable supports include glass, plastic, metal, polymer gels, and the like, and may take the form of beads, wells, dipstics, membranes, and the like.

The term "label" as used herein refers to any atom or moiety which can be used to provide a detectable (preferably quantifiable) signal, and which can be attached to a polynucleotide or polypeptide.

As used herein, the term "label probe" refers to an oligomer which is comprised of targeting polynucleotide sequence, which is complementary to a target sequence to be detected in the analyte polynucleotide. This complementary region is of sufficient length and complementarity to the target sequence to afford a duplex comprised of the "label probe" and the "target sequence" to be detected by the label. The oligomer is coupled to a

label eith r directly, or indirectly via a set of ligand molecules with high specificity for each other. Sets of ligand molecules with high specificity are described supra., and also includes multimers.

The term "multimer", as used herein, refers to linear or branched polymers of the same repeating single-stranded polynucleotide unit or different single-stranded polynucleotide units. At least one of the units has a sequence, length, and composition that permits it to hybridize specifically to a first single-stranded nucleotide sequence of interest, typically an analyte or an oligomer (e.g., a label probe) bound to an analyte. In order to achieve such specificity and stability, this unit will normally be at least about 15 nucleotides in length, typically no more than about 50 nucleotides in length, and preferably about 30 nucleotides in length; moreover, the content of Gs and Cs will normally be at least about 40%, and at most about 60%. In addition to such unit(s), the multimer includes a multiplicity of units that are capable of hybridizing specifically and stably to a second single-stranded nucleotide of interest, typically a labeled polynucleotide or another multimer. These units are generally about the same size and composition as the multimers discussed above. When a multimer is designed to be hybridized to another multimer, the first and second oligonucleotide units are heterogeneous (different), and do not hybridize with each other under the conditions of the selected assay. Thus, multimers may be label probes, or may be ligands which couple the label to the probe.

As used herein, the term "viral RNA", which includes HCV RNA, refers to RNA from the viral genome, fragments thereof, transcripts thereof, and mutant sequences derived therefrom.

As used herein, a "biological sample" refers to a sample of tissue or fluid isolated from an individual, including but not limited to, for example, plasma, serum,

spinal fluid, lymph fluid, the external sections of the skin, respiratory, intestinal, and genitourinary tracts, tears, saliva, milk, blood cells, tumors, organs, and also samples of in vitro cell culture constitu nts (including but not limited to conditioned medium resulting from the growth of cells in cell culture medium, putatively virally infected cells, recombinant cells, and cell components).

Description of the Invention

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See e.g., Maniatis, Fitsch & Sambrook, MOLECULAR CLONING; A LABORATORY MANUAL (1982); DNA CLONING, VOLUMES I AND II (D.N Glover ed. 1985); OLIGONUCLEOTIDE SYNTHESIS (M.J. Gait ed, 1984); NUCLEIC ACID HYBRIDIZATION (B.D. Hames & S.J. Higgins eds. 1984); the series, METHODS IN ENZYMOLOGY (Academic Press, Inc.), particularly Vol. 154 and Vol. 155 (Wu and Grossman, and Wu, eds., respectively). All patents, patent applications, and publications mentioned herein, both supra and infra, are hereby incorporated herein by reference.

The useful materials and processes of the present invention are made possible by the identification of HCV as the etiologic agent of BB-NANBV, and by the provision of a family of nucleotide sequences isolated from cDNA libraries which contain HCV cDNA sequences. These cDNA libraries were derived from nucleic acid sequences present in the plasma of an HCV-infected chimpanzee. The construction of one of these libraries, the "c" library (ATCC No. 40394), is described in E.P.O. Publication No. 318,216.

Utilizing the above-described HCV cDNA sequences, as well as that described herein, oligomers can

be constructed which are useful as reagents for d tecting viral polynucl otides in biological samples. For example, from th s quenc s it is possible to synthesiz DNA oligomers of about 8-10 nucleotides, or larger, which are useful as hybridization probes to detect the presence of HCV RNA in, for example, donated blood, blood fractions, sera of subjects suspected of harboring the virus, or cell culture systems in which the virus is replicating. In addition, the novel oligomers described herein enable further characterization of the HCV genome. Polynucleotide probes and primers derived from these sequences may be used to amplify sequences present in cDNA libraries, and/or to screen cDNA libraries for additional overlapping cDNA sequences, which, in turn, may be used to obtain more overlapping sequences. As indicated infra. and in E.P.O. Publication No. 318,216, the genome of HCV appears to be RNA comprised primarily of a large open reading frame (ORF) which encodes a large polyprotein.

In addition to the above, the information provided infra allows the identification of additional HCV strains or isolates. The isolation and characterization of the additional HCV strains or isolates may be accomplished utilizing techniques known to those of skill in the art, for example, by isolating the nucleic acids from body components which contain viral particles and/or viral RNA, creating cDNA libraries using the oligomers described infra., for screening the libraries for clones containing HCV cDNA sequences described infra., and comparing the HCV cDNAs from the new isolates with the cDNAs described in E.P.O. Publication No. 318,216 and infra. Strains or isolates which fit within the parameters of HCV, as described in the Definitions section, supra., are readily identifiable. Other methods for identifying HCV strains will be obvious to those of skill in the art, based upon the information provided herein.

Isolation of the HCV cDNA Sequences

Th oligomers of the invention contain regions which form hybrid duplex structures with targeted sequences in HCV polynucleotides. The HCV polynucleotid hybridizing regions of the oligomers may be ascertained from the HCV cDNA sequence(s) provided herein, and described in E.P.O. Publication No. 318,216. A composite of HCV cDNA from HCV1, a prototypic HCV, is shown in Fig. 18. The composite sequence is based upon sequence information derived from a number of HCV cDNA clones, which were isolated from a number of HCV cDNA libraries, including the "c" library present in lambda gt11 (ATCC No. 40394), and from human serum. The HCV cDNA clones were isolated by methods described in E.P.O. Publication No. 318,216. Briefly, the majority of clones which were isolated contained sequences from the HCV cDNA "c" library which was constructed using pooled serum from a chimpanzee with chronic HCV infection and containing a high titer of the virus, i.e., at least $10^6$ chimp infectious doses/ml (CID/ml). The pooled serum was used to isolate viral particles; nucleic acids isolated from these particles was used as the template in the construction of cDNA libraries to the viral genome. The initial clone, 5-1-1, was obtained by screening the "c" library with serum from infected individuals. After the isolation of the initial clone, the remainder of the sequence was obtained by screening with synthetic polynucleotide probes, the sequences of which were derived from the 5'-region and the 3'-region of the known HCV cDNA sequence(s).

The description of the methods to retrieve the cDNA sequences is mostly of historical interest. The resultant sequences (and their complements) are provided herein, and the sequences, or any portion thereof, could be prepared using synthetic methods, or by a combination of synthetic methods with retrieval of partial sequences

using m thods similar to those describ d in E.P.O. Publication No. 318,216.

Oligom r Prob s and Prim rs

Using as a basis the HCV genome (as illustrated in Fig. 18), and/or preferably conserved regions of the HCV genome, oligomers of approximately 8 nucleotides or more can be prepared which hybridize with the positive strand(s) of HCV RNA or its complement, as well as to HCV cDNAs. These oligomers can serve as probes for the detection (including isolation and/or labeling) of polynucleotides which contain HCV nucleotide sequences, and/or as primers for the transcription and/or replication of targeted HCV sequences. The oligomers contain a targeting polynucleotide sequence, which is comprised of nucleotides which are complementary to a target HCV nucleotide sequence; the sequence is of sufficient length and complementarity with the HCV sequence to form a duplex which has sufficient stability for the purpose intended. For example, if the purpose is the isolation, via immobilization, of an analyte containing a target HCV sequence, the oligomers would contain a polynucleotide region which is of sufficient length and complementarity to the targeted HCV sequence to afford sufficient duplex stability to immobilize the analyte on a solid surface, via its binding to the oligomers, under the isolation conditions. For example, also, if the oligomers are to serve as primers for the transcription and/or replication of target HCV sequences in an analyte polynucleotide, the oligomers would contain a polynucleotide region of sufficient length and complementarity to the targeted HCV sequence to allow the polymerizing agent to continue replication from the primers which are in stable duplex form with the target sequence, under the polymerizing conditions. For example, also, if the oligomers are to be used as label probes, or are to bind to multimers, the

targeting polynucleotide region would be of suffici nt l ngth and complementarity to form stable hybrid duplex structures with the label probes and/or multim rs to allow detection of th duplex. The oligom rs may contain a minimum of about 4 contiguous nucleotides which are complementary to targeted HCV sequence; usually the oligomers will contain a minimum of about 8 continguous nucleotides which are complementary to the targeted HCV sequence, and preferably will contain a minimum of about 14 contiguous nucleotides which are complementary to the targeted HCV sequence.

Suitable HCV nucleotide targeting sequences may be comprised of nucleotides which are complementary nucleotides selected from the following HCV cDNA nucleotides, which are shown in Fig. 18, ($nn_x$ - $nn_y$ denotes from about nucleotide number x to about nucleotide number y)):

$nn_{-340}$ - $nn_{-330}$; $nn_{-330}$ - $nn_{-320}$; $nn_{-320}$ - $nn_{-310}$;
$nn_{-310}$ - $nn_{-300}$; $nn_{-300}$ - $nn_{-290}$; $nn_{-290}$ - $nn_{-280}$;
$nn_{-280}$ - $nn_{-270}$; $nn_{-270}$ - $nn_{-260}$; $nn_{-260}$ - $nn_{-250}$;
$nn_{-250}$ - $nn_{-240}$; $nn_{-240}$ - $nn_{-230}$; $nn_{-230}$ - $nn_{-220}$;
$nn_{-220}$ - $nn_{-210}$; $nn_{-210}$ - $nn_{-200}$; $nn_{-200}$ - $nn_{-190}$;
$nn_{-190}$ - $nn_{-180}$; $nn_{-180}$ - $nn_{-170}$; $nn_{-170}$ - $nn_{-160}$;
$nn_{-160}$ - $nn_{-150}$; $nn_{-150}$ - $nn_{-140}$; $nn_{-140}$ - $nn_{-130}$;
$nn_{-130}$ - $nn_{-120}$; $nn_{-120}$ - $nn_{-110}$; $nn_{-110}$ - $nn_{-100}$;
$nn_{-100}$ - $nn_{-90}$; $nn_{-90}$ - $nn_{-80}$; $nn_{-80}$ - $nn_{-70}$;
$nn_{-70}$ - $nn_{-60}$; $nn_{-60}$ - $nn_{-50}$; $nn_{-50}$ - $nn_{-40}$;
$nn_{-40}$ - $nn_{-30}$; $nn_{-30}$ - $nn_{-20}$; $nn_{-20}$ - $nn_{-10}$;
$nn_{-10}$ - $nn_{1}$; $nn_{1}$ - $nn_{10}$; $nn_{10}$ - $nn_{20}$; $nn_{20}$ - $nn_{30}$;
$nn_{30}$ - $nn_{40}$; $nn_{40}$ - $nn_{50}$; $nn_{50}$ - $nn_{60}$; $nn_{60}$ - $nn_{70}$;
$nn_{70}$ - $nn_{80}$; $nn_{80}$ - $nn_{90}$; $nn_{90}$ - $nn_{100}$; $nn_{100}$ - $nn_{110}$;
$nn_{110}$ - $nn_{120}$; $nn_{120}$ - $nn_{130}$; $nn_{130}$ - $nn_{140}$;
$nn_{140}$ - $nn_{150}$; $nn_{150}$ - $nn_{160}$; $nn_{160}$ - $nn_{170}$;
$nn_{170}$ - $nn_{180}$; $nn_{180}$ - $nn_{190}$; $nn_{190}$ - $nn_{200}$;
$nn_{200}$ - $nn_{210}$; $nn_{210}$ - $nn_{220}$; $nn_{220}$ - $nn_{230}$;

$nn_{230}$ - $nn_{240}$; $nn_{240}$ - $nn_{250}$; $nn_{250}$ - $nn_{260}$; $nn_{260}$ - $nn_{270}$; $nn_{270}$ - $nn_{280}$; $nn_{280}$ - $nn_{290}$; $nn_{290}$ - $nn_{300}$; $nn_{300}$ - $nn_{310}$; $nn_{310}$ - $nn_{320}$; $nn_{320}$ - $nn_{330}$; $nn_{330}$ - $nn_{340}$; $nn_{340}$ - $nn_{350}$; $nn_{350}$ - $nn_{360}$; $nn_{360}$ - $nn_{370}$; $nn_{370}$ - $nn_{380}$; $nn_{380}$ - $nn_{390}$; $nn_{390}$ - $nn_{400}$; $nn_{400}$ - $nn_{410}$; $nn_{410}$ - $nn_{420}$; $nn_{420}$ - $nn_{430}$; $nn_{430}$ - $nn_{440}$; $nn_{440}$ - $nn_{450}$; $nn_{450}$ - $nn_{460}$; $nn_{460}$ - $nn_{470}$; $nn_{470}$ - $nn_{480}$; $nn_{480}$ - $nn_{490}$; $nn_{490}$ - $nn_{500}$; $nn_{500}$ - $nn_{510}$; $nn_{510}$ - $nn_{520}$; $nn_{520}$ - $nn_{530}$; $nn_{530}$ - $nn_{540}$; $nn_{540}$ - $nn_{550}$; $nn_{550}$ - $nn_{560}$; $nn_{560}$ - $nn_{570}$; $nn_{570}$ - $nn_{580}$; $nn_{580}$ - $nn_{590}$; $nn_{590}$ - $nn_{600}$; $nn_{600}$ - $nn_{610}$; $nn_{610}$ - $nn_{620}$; $nn_{620}$ - $nn_{630}$; $nn_{630}$ - $nn_{640}$; $nn_{640}$ - $nn_{650}$; $nn_{650}$ - $nn_{660}$; $nn_{660}$ - $nn_{670}$; $nn_{670}$ - $nn_{680}$; $nn_{680}$ - $nn_{690}$; $nn_{690}$ - $nn_{700}$; $nn_{700}$ - $nn_{710}$; $nn_{710}$ - $nn_{720}$; $nn_{720}$ - $nn_{730}$; $nn_{730}$ - $nn_{740}$; $nn_{740}$ - $nn_{750}$; $nn_{750}$ - $nn_{760}$; $nn_{760}$ - $nn_{770}$; $nn_{770}$ - $nn_{780}$; $nn_{780}$ - $nn_{790}$; $nn_{790}$ - $nn_{800}$; $nn_{800}$ - $nn_{810}$; $nn_{810}$ - $nn_{820}$; $nn_{820}$ - $nn_{830}$; $nn_{830}$ - $nn_{840}$; $nn_{840}$ - $nn_{850}$; $nn_{850}$ - $nn_{860}$; $nn_{860}$ - $nn_{870}$; $nn_{870}$ - $nn_{880}$; $nn_{880}$ - $nn_{890}$; $nn_{890}$ - $nn_{900}$; $nn_{900}$ - $nn_{910}$; $nn_{910}$ - $nn_{920}$; $nn_{920}$ - $nn_{930}$; $nn_{930}$ - $nn_{940}$; $nn_{940}$ - $nn_{950}$; $nn_{950}$ - $nn_{960}$; $nn_{960}$ - $nn_{970}$; $nn_{970}$ - $nn_{980}$; $nn_{980}$ - $nn_{990}$; $nn_{990}$ - $nn_{1000}$; $nn_{1000}$ - $nn_{1010}$; $nn_{1010}$ - $nn_{1020}$; $nn_{1020}$ - $nn_{1030}$; $nn_{1030}$ - $nn_{1040}$; $nn_{1040}$ - $nn_{1050}$; $nn_{1050}$ - $nn_{1060}$; $nn_{1060}$ - $nn_{1070}$; $nn_{1070}$ - $nn_{1080}$; $nn_{1080}$ - $nn_{1090}$; $nn_{1090}$ - $nn_{1100}$; $nn_{1100}$ - $nn_{1110}$; $nn_{1110}$ - $nn_{1120}$; $nn_{1120}$ - $nn_{1130}$; $nn_{1130}$ - $nn_{1140}$; $nn_{1140}$ - $nn_{1150}$; $nn_{1150}$ - $nn_{1160}$; $nn_{1160}$ - $nn_{1170}$; $nn_{1170}$ - $nn_{1180}$; $nn_{1180}$ - $nn_{1190}$; $nn_{1190}$ - $nn_{1200}$; $nn_{1200}$ - $nn_{1210}$; $nn_{1210}$ - $nn_{1220}$; $nn_{1220}$ - $nn_{1230}$; $nn_{1230}$ - $nn_{1240}$; $nn_{1240}$ - $nn_{1250}$; $nn_{1250}$ - $nn_{1260}$; $nn_{1260}$ - $nn_{1270}$; $nn_{1270}$ - $nn_{1280}$; $nn_{1280}$ - $nn_{1290}$; $nn_{1290}$ - $nn_{1300}$; $nn_{1300}$ - $nn_{1310}$;

$nn_{1310}$ - $nn_{1320}$; $nn_{1320}$ - $nn_{1330}$; $nn_{1330}$ - $nn_{1340}$; $nn_{1340}$ - $nn_{1350}$; $nn_{1350}$ - $nn_{1360}$; $nn_{1360}$ - $nn_{1370}$; $nn_{1370}$ - $nn_{1380}$; $nn_{1380}$ - $nn_{1390}$; $nn_{1390}$ - $nn_{1400}$; $nn_{1400}$ - $nn_{1410}$; $nn_{1410}$ - $nn_{1420}$; $nn_{1420}$ - $nn_{1430}$; $nn_{1430}$ - $nn_{1440}$; $nn_{1440}$ - $nn_{1450}$; $nn_{1450}$ - $nn_{1460}$; $nn_{1460}$ - $nn_{1470}$; $nn_{1470}$ - $nn_{1480}$; $nn_{1480}$ - $nn_{1490}$; $nn_{1490}$ - $nn_{1500}$; $nn_{1500}$ - $nn_{1510}$; $nn_{1510}$ - $nn_{1520}$; $nn_{1520}$ - $nn_{1530}$; $nn_{1530}$ - $nn_{1540}$; $nn_{1540}$ - $nn_{1550}$; $nn_{1550}$ - $nn_{1560}$; $nn_{1560}$ - $nn_{1570}$; $nn_{1570}$ - $nn_{1580}$; $nn_{1580}$ - $nn_{1590}$; $nn_{1590}$ - $nn_{1600}$; $nn_{1600}$ - $nn_{1610}$; $nn_{1610}$ - $nn_{1620}$; $nn_{1620}$ - $nn_{1630}$; $nn_{1630}$ - $nn_{1640}$; $nn_{1640}$ - $nn_{1650}$; $nn_{1650}$ - $nn_{1660}$; $nn_{1660}$ - $nn_{1670}$; $nn_{1670}$ - $nn_{1680}$; $nn_{1680}$ - $nn_{1690}$; $nn_{1690}$ - $nn_{1700}$; $nn_{1700}$ - $nn_{1710}$; $nn_{1710}$ - $nn_{1720}$; $nn_{1720}$ - $nn_{1730}$; $nn_{1730}$ - $nn_{1740}$; $nn_{1740}$ - $nn_{1750}$; $nn_{1750}$ - $nn_{1760}$; $nn_{1760}$ - $nn_{1770}$; $nn_{1770}$ - $nn_{1780}$; $nn_{1780}$ - $nn_{1790}$; $nn_{1790}$ - $nn_{1800}$; $nn_{1800}$ - $nn_{1810}$; $nn_{1810}$ - $nn_{1820}$; $nn_{1820}$ - $nn_{1830}$; $nn_{1830}$ - $nn_{1840}$; $nn_{1840}$ - $nn_{1850}$; $nn_{1850}$ - $nn_{1860}$; $nn_{1860}$ - $nn_{1870}$; $nn_{1870}$ - $nn_{1880}$; $nn_{1880}$ - $nn_{1890}$; $nn_{1890}$ - $nn_{1900}$; $nn_{1900}$ - $nn_{1910}$; $nn_{1910}$ - $nn_{1920}$; $nn_{1920}$ - $nn_{1930}$; $nn_{1930}$ - $nn_{1940}$; $nn_{1940}$ - $nn_{1950}$; $nn_{1950}$ - $nn_{1960}$; $nn_{1960}$ - $nn_{1970}$; $nn_{1970}$ - $nn_{1980}$; $nn_{1980}$ - $nn_{1990}$; $nn_{1990}$ - $nn_{2000}$; $nn_{2000}$ - $nn_{2010}$; $nn_{2010}$ - $nn_{2020}$; $nn_{2020}$ - $nn_{2030}$; $nn_{2030}$ - $nn_{2040}$; $nn_{2040}$ - $nn_{2050}$; $nn_{2050}$ - $nn_{2060}$; $nn_{2060}$ - $nn_{2070}$; $nn_{2070}$ - $nn_{2080}$; $nn_{2080}$ - $nn_{2090}$; $nn_{2090}$ - $nn_{2100}$; $nn_{2100}$ - $nn_{2110}$; $nn_{2110}$ - $nn_{2120}$; $nn_{2120}$ - $nn_{2130}$; $nn_{2130}$ - $nn_{2140}$; $nn_{2140}$ - $nn_{2150}$; $nn_{2150}$ - $nn_{2160}$; $nn_{2160}$ - $nn_{2170}$; $nn_{2170}$ - $nn_{2180}$; $nn_{2180}$ - $nn_{2190}$; $nn_{2190}$ - $nn_{2200}$; $nn_{2200}$ - $nn_{2210}$; $nn_{2210}$ - $nn_{2220}$; $nn_{2220}$ - $nn_{2230}$; $nn_{2230}$ - $nn_{2240}$; $nn_{2240}$ - $nn_{2250}$; $nn_{2250}$ - $nn_{2260}$; $nn_{2260}$ - $nn_{2270}$; $nn_{2270}$ - $nn_{2280}$; $nn_{2280}$ - $nn_{2290}$; $nn_{2290}$ - $nn_{2300}$; $nn_{2300}$ - $nn_{2310}$; $nn_{2310}$ - $nn_{2320}$; $nn_{2320}$ - $nn_{2330}$; $nn_{2330}$ - $nn_{2340}$; $nn_{2340}$ - $nn_{2350}$; $nn_{2350}$ - $nn_{2360}$; $nn_{2360}$ - $nn_{2370}$; $nn_{2370}$ - $nn_{2380}$; $nn_{2380}$ - $nn_{2390}$;

$nn_{2390}$ - $nn_{2400}$; $nn_{2400}$ - $nn_{2410}$; $nn_{2410}$ - $nn_{2420}$; $nn_{2420}$ - $nn_{2430}$; $nn_{2430}$ - $nn_{2440}$; $nn_{2440}$ - $nn_{2450}$; $nn_{2450}$ - $nn_{2460}$; $nn_{2460}$ - $nn_{2470}$; $nn_{2470}$ - $nn_{2480}$; $nn_{2480}$ - $nn_{2490}$; $nn_{2490}$ - $nn_{2500}$; $nn_{2500}$ - $nn_{2510}$; $nn_{2510}$ - $nn_{2520}$; $nn_{2520}$ - $nn_{2530}$; $nn_{2530}$ - $nn_{2540}$; $nn_{2540}$ - $nn_{2550}$; $nn_{2550}$ - $nn_{2560}$; $nn_{2560}$ - $nn_{2570}$; $nn_{2570}$ - $nn_{2580}$; $nn_{2580}$ - $nn_{2590}$; $nn_{2590}$ - $nn_{2600}$; $nn_{2600}$ - $nn_{2610}$; $nn_{2610}$ - $nn_{2620}$; $nn_{2620}$ - $nn_{2630}$; $nn_{2630}$ - $nn_{2640}$; $nn_{2640}$ - $nn_{2650}$; $nn_{2650}$ - $nn_{2660}$; $nn_{2660}$ - $nn_{2670}$; $nn_{2670}$ - $nn_{2680}$; $nn_{2680}$ - $nn_{2690}$; $nn_{2690}$ - $nn_{2700}$; $nn_{2700}$ - $nn_{2710}$; $nn_{2710}$ - $nn_{2720}$; $nn_{2720}$ - $nn_{2730}$; $nn_{2730}$ - $nn_{2740}$; $nn_{2740}$ - $nn_{2750}$; $nn_{2750}$ - $nn_{2760}$; $nn_{2760}$ - $nn_{2770}$; $nn_{2770}$ - $nn_{2780}$; $nn_{2780}$ - $nn_{2790}$; $nn_{2790}$ - $nn_{2800}$; $nn_{2800}$ - $nn_{2810}$; $nn_{2810}$ - $nn_{2820}$; $nn_{2820}$ - $nn_{2830}$; $nn_{2830}$ - $nn_{2840}$; $nn_{2840}$ - $nn_{2850}$; $nn_{2850}$ - $nn_{2860}$; $nn_{2860}$ - $nn_{2870}$; $nn_{2870}$ - $nn_{2880}$; $nn_{2880}$ - $nn_{2890}$; $nn_{2890}$ - $nn_{2900}$; $nn_{2900}$ - $nn_{2910}$; $nn_{2910}$ - $nn_{2920}$; $nn_{2920}$ - $nn_{2930}$; $nn_{2930}$ - $nn_{2940}$; $nn_{2940}$ - $nn_{2950}$; $nn_{2950}$ - $nn_{2960}$; $nn_{2960}$ - $nn_{2970}$; $nn_{2970}$ - $nn_{2980}$; $nn_{2980}$ - $nn_{2990}$; $nn_{2990}$ - $nn_{3000}$; $nn_{3000}$ - $nn_{3010}$; $nn_{3010}$ - $nn_{3020}$; $nn_{3020}$ - $nn_{3030}$; $nn_{3030}$ - $nn_{3040}$; $nn_{3040}$ - $nn_{3050}$; $nn_{3050}$ - $nn_{3060}$; $nn_{3060}$ - $nn_{3070}$; $nn_{3070}$ - $nn_{3080}$; $nn_{3080}$ - $nn_{3090}$; $nn_{3090}$ - $nn_{3100}$; $nn_{3100}$ - $nn_{3110}$; $nn_{3110}$ - $nn_{3120}$; $nn_{3120}$ - $nn_{3130}$; $nn_{3130}$ - $nn_{3140}$; $nn_{3140}$ - $nn_{3150}$; $nn_{3150}$ - $nn_{3160}$; $nn_{3160}$ - $nn_{3170}$; $nn_{3170}$ - $nn_{3180}$; $nn_{3180}$ - $nn_{3190}$; $nn_{3190}$ - $nn_{3200}$; $nn_{3200}$ - $nn_{3210}$; $nn_{3210}$ - $nn_{3220}$; $nn_{3220}$ - $nn_{3230}$; $nn_{3230}$ - $nn_{3240}$; $nn_{3240}$ - $nn_{3250}$; $nn_{3250}$ - $nn_{3260}$; $nn_{3260}$ - $nn_{3270}$; $nn_{3270}$ - $nn_{3280}$; $nn_{3280}$ - $nn_{3290}$; $nn_{3290}$ - $nn_{3300}$; $nn_{3300}$ - $nn_{3310}$; $nn_{3310}$ - $nn_{3320}$; $nn_{3320}$ - $nn_{3330}$; $nn_{3330}$ - $nn_{3340}$; $nn_{3340}$ - $nn_{3350}$; $nn_{3350}$ - $nn_{3360}$; $nn_{3360}$ - $nn_{3370}$; $nn_{3370}$ - $nn_{3380}$; $nn_{3380}$ - $nn_{3390}$; $nn_{3390}$ - $nn_{3400}$; $nn_{3400}$ - $nn_{3410}$; $nn_{3410}$ - $nn_{3420}$; $nn_{3420}$ - $nn_{3430}$; $nn_{3430}$ - $nn_{3440}$; $nn_{3440}$ - $nn_{3450}$; $nn_{3450}$ - $nn_{3460}$; $nn_{3460}$ - $nn_{3470}$;

$nn_{3470}$ - $nn_{3480}$; $nn_{3480}$ - $nn_{3490}$; $nn_{3490}$ - $nn_{3500}$;
$nn_{3500}$ - $nn_{3510}$; $nn_{3510}$ - $nn_{3520}$; $nn_{3520}$ - $nn_{3530}$;
$nn_{3530}$ - $nn_{3540}$; $nn_{3540}$ - $nn_{3550}$; $nn_{3550}$ - $nn_{3560}$;
$nn_{3560}$ - $nn_{3570}$; $nn_{3570}$ - $nn_{3580}$; $nn_{3580}$ - $nn_{3590}$;
$nn_{3590}$ - $nn_{3600}$; $nn_{3600}$ - $nn_{3610}$; $nn_{3610}$ - $nn_{3620}$;
$nn_{3620}$ - $nn_{3630}$; $nn_{3630}$ - $nn_{3640}$; $nn_{3640}$ - $nn_{3650}$;
$nn_{3650}$ - $nn_{3660}$; $nn_{3660}$ - $nn_{3670}$; $nn_{3670}$ - $nn_{3680}$;
$nn_{3680}$ - $nn_{3690}$; $nn_{3690}$ - $nn_{3700}$; $nn_{3700}$ - $nn_{3710}$;
$nn_{3710}$ - $nn_{3720}$; $nn_{3720}$ - $nn_{3730}$; $nn_{3730}$ - $nn_{3740}$;
$nn_{3740}$ - $nn_{3750}$; $nn_{3750}$ - $nn_{3760}$; $nn_{3760}$ - $nn_{3770}$;
$nn_{3770}$ - $nn_{3780}$; $nn_{3780}$ - $nn_{3790}$; $nn_{3790}$ - $nn_{3800}$;
$nn_{3800}$ - $nn_{3810}$; $nn_{3810}$ - $nn_{3820}$; $nn_{3820}$ - $nn_{3830}$;
$nn_{3830}$ - $nn_{3840}$; $nn_{3840}$ - $nn_{3850}$; $nn_{3850}$ - $nn_{3860}$;
$nn_{3860}$ - $nn_{3870}$; $nn_{3870}$ - $nn_{3880}$; $nn_{3880}$ - $nn_{3890}$;
$nn_{3890}$ - $nn_{3900}$; $nn_{3900}$ - $nn_{3910}$; $nn_{3910}$ - $nn_{3920}$;
$nn_{3920}$ - $nn_{3930}$; $nn_{3930}$ - $nn_{3940}$; $nn_{3940}$ - $nn_{3950}$;
$nn_{3950}$ - $nn_{3960}$; $nn_{3960}$ - $nn_{3970}$; $nn_{3970}$ - $nn_{3980}$;
$nn_{3980}$ - $nn_{3990}$; $nn_{3990}$ - $nn_{4000}$; $nn_{4000}$ - $nn_{4010}$;
$nn_{4010}$ - $nn_{4020}$; $nn_{4020}$ - $nn_{4030}$; $nn_{4030}$ - $nn_{4040}$;
$nn_{4040}$ - $nn_{4050}$; $nn_{4050}$ - $nn_{4060}$; $nn_{4060}$ - $nn_{4070}$;
$nn_{4070}$ - $nn_{4080}$; $nn_{4080}$ - $nn_{4090}$; $nn_{4090}$ - $nn_{4100}$;
$nn_{4100}$ - $nn_{4110}$; $nn_{4110}$ - $nn_{4120}$; $nn_{4120}$ - $nn_{4130}$;
$nn_{4130}$ - $nn_{4140}$; $nn_{4140}$ - $nn_{4150}$; $nn_{4150}$ - $nn_{4160}$;
$nn_{4160}$ - $nn_{4170}$; $nn_{4170}$ - $nn_{4180}$; $nn_{4180}$ - $nn_{4190}$;
$nn_{4190}$ - $nn_{4200}$; $nn_{4200}$ - $nn_{4210}$; $nn_{4210}$ - $nn_{4220}$;
$nn_{4220}$ - $nn_{4230}$; $nn_{4230}$ - $nn_{4240}$; $nn_{4240}$ - $nn_{4250}$;
$nn_{4250}$ - $nn_{4260}$; $nn_{4260}$ - $nn_{4270}$; $nn_{4270}$ - $nn_{4280}$;
$nn_{4280}$ - $nn_{4290}$; $nn_{4290}$ - $nn_{4300}$; $nn_{4300}$ - $nn_{4310}$;
$nn_{4310}$ - $nn_{4320}$; $nn_{4320}$ - $nn_{4330}$; $nn_{4330}$ - $nn_{4340}$;
$nn_{4340}$ - $nn_{4350}$; $nn_{4350}$ - $nn_{4360}$; $nn_{4360}$ - $nn_{4370}$;
$nn_{4370}$ - $nn_{4380}$; $nn_{4380}$ - $nn_{4390}$; $nn_{4390}$ - $nn_{4400}$;
$nn_{4400}$ - $nn_{4410}$; $nn_{4410}$ - $nn_{4420}$; $nn_{4420}$ - $nn_{4430}$;
$nn_{4430}$ - $nn_{4440}$; $nn_{4440}$ - $nn_{4450}$; $nn_{4450}$ - $nn_{4460}$;
$nn_{4460}$ - $nn_{4470}$; $nn_{4470}$ - $nn_{4480}$; $nn_{4480}$ - $nn_{4490}$;
$nn_{4490}$ - $nn_{4500}$; $nn_{4500}$ - $nn_{4510}$; $nn_{4510}$ - $nn_{4520}$;
$nn_{4520}$ - $nn_{4530}$; $nn_{4530}$ - $nn_{4540}$; $nn_{4540}$ - $nn_{4550}$;

$nn_{4550}$ - $nn_{4560}$; $nn_{4560}$ - $nn_{4570}$; $nn_{4570}$ - $nn_{4580}$; $nn_{4580}$ - $nn_{4590}$; $nn_{4590}$ - $nn_{4600}$; $nn_{4600}$ - $nn_{4610}$; $nn_{4610}$ - $nn_{4620}$; $nn_{4620}$ - $nn_{4630}$; $nn_{4630}$ - $nn_{4640}$; $nn_{4640}$ - $nn_{4650}$; $nn_{4650}$ - $nn_{4660}$; $nn_{4660}$ - $nn_{4670}$; $nn_{4670}$ - $nn_{4680}$; $nn_{4680}$ - $nn_{4690}$; $nn_{4690}$ - $nn_{4700}$; $nn_{4700}$ - $nn_{4710}$; $nn_{4710}$ - $nn_{4720}$; $nn_{4720}$ - $nn_{4730}$; $nn_{4730}$ - $nn_{4740}$; $nn_{4740}$ - $nn_{4750}$; $nn_{4750}$ - $nn_{4760}$; $nn_{4760}$ - $nn_{4770}$; $nn_{4770}$ - $nn_{4780}$; $nn_{4780}$ - $nn_{4790}$; $nn_{4790}$ - $nn_{4800}$; $nn_{4800}$ - $nn_{4810}$; $nn_{4810}$ - $nn_{4820}$; $nn_{4820}$ - $nn_{4830}$; $nn_{4830}$ - $nn_{4840}$; $nn_{4840}$ - $nn_{4850}$; $nn_{4850}$ - $nn_{4860}$; $nn_{4860}$ - $nn_{4870}$; $nn_{4870}$ - $nn_{4880}$; $nn_{4880}$ - $nn_{4890}$; $nn_{4890}$ - $nn_{4900}$; $nn_{4900}$ - $nn_{4910}$; $nn_{4910}$ - $nn_{4920}$; $nn_{4920}$ - $nn_{4930}$; $nn_{4930}$ - $nn_{4940}$; $nn_{4940}$ - $nn_{4950}$; $nn_{4950}$ - $nn_{4960}$; $nn_{4960}$ - $nn_{4970}$; $nn_{4970}$ - $nn_{4980}$; $nn_{4980}$ - $nn_{4990}$; $nn_{4990}$ - $nn_{5000}$; $nn_{5000}$ - $nn_{5010}$; $nn_{5010}$ - $nn_{5020}$; $nn_{5020}$ - $nn_{5030}$; $nn_{5030}$ - $nn_{5040}$; $nn_{5040}$ - $nn_{5050}$; $nn_{5050}$ - $nn_{5060}$; $nn_{5060}$ - $nn_{5070}$; $nn_{5070}$ - $nn_{5080}$; $nn_{5080}$ - $nn_{5090}$; $nn_{5090}$ - $nn_{5100}$; $nn_{5100}$ - $nn_{5110}$; $nn_{5110}$ - $nn_{5120}$; $nn_{5120}$ - $nn_{5130}$; $nn_{5130}$ - $nn_{5140}$; $nn_{5140}$ - $nn_{5150}$; $nn_{5150}$ - $nn_{5160}$; $nn_{5160}$ - $nn_{5170}$; $nn_{5170}$ - $nn_{5180}$; $nn_{5180}$ - $nn_{5190}$; $nn_{5190}$ - $nn_{5200}$; $nn_{5200}$ - $nn_{5210}$; $nn_{5210}$ - $nn_{5220}$; $nn_{5220}$ - $nn_{5230}$; $nn_{5230}$ - $nn_{5240}$; $nn_{5240}$ - $nn_{5250}$; $nn_{5250}$ - $nn_{5260}$; $nn_{5260}$ - $nn_{5270}$; $nn_{5270}$ - $nn_{5280}$; $nn_{5280}$ - $nn_{5290}$; $nn_{5290}$ - $nn_{5300}$; $nn_{5300}$ - $nn_{5310}$; $nn_{5310}$ - $nn_{5320}$; $nn_{5320}$ - $nn_{5330}$; $nn_{5330}$ - $nn_{5340}$; $nn_{5340}$ - $nn_{5350}$; $nn_{5350}$ - $nn_{5360}$; $nn_{5360}$ - $nn_{5370}$; $nn_{5370}$ - $nn_{5380}$; $nn_{5380}$ - $nn_{5390}$; $nn_{5390}$ - $nn_{5400}$; $nn_{5400}$ - $nn_{5410}$; $nn_{5410}$ - $nn_{5420}$; $nn_{5420}$ - $nn_{5430}$; $nn_{5430}$ - $nn_{5440}$; $nn_{5440}$ - $nn_{5450}$; $nn_{5450}$ - $nn_{5460}$; $nn_{5460}$ - $nn_{5470}$; $nn_{5470}$ - $nn_{5480}$; $nn_{5480}$ - $nn_{5490}$; $nn_{5490}$ - $nn_{5500}$; $nn_{5500}$ - $nn_{5510}$; $nn_{5510}$ - $nn_{5520}$; $nn_{5520}$ - $nn_{5530}$; $nn_{5530}$ - $nn_{5540}$; $nn_{5540}$ - $nn_{5550}$; $nn_{5550}$ - $nn_{5560}$; $nn_{5560}$ - $nn_{5570}$; $nn_{5570}$ - $nn_{5580}$; $nn_{5580}$ - $nn_{5590}$; $nn_{5590}$ - $nn_{5600}$; $nn_{5600}$ - $nn_{5610}$; $nn_{5610}$ - $nn_{5620}$; $nn_{5620}$ - $nn_{5630}$;

$nn_{5630}$ - $nn_{5640}$; $nn_{5640}$ - $nn_{5650}$; $nn_{5650}$ - $nn_{5660}$; $nn_{5660}$ - $nn_{5670}$; $nn_{5670}$ - $nn_{5680}$; $nn_{5680}$ - $nn_{5690}$; $nn_{5690}$ - $nn_{5700}$; $nn_{5700}$ - $nn_{5710}$; $nn_{5710}$ - $nn_{5720}$; $nn_{5720}$ - $nn_{5730}$; $nn_{5730}$ - $nn_{5740}$; $nn_{5740}$ - $nn_{5750}$; $nn_{5750}$ - $nn_{5760}$; $nn_{5760}$ - $nn_{5770}$; $nn_{5770}$ - $nn_{5780}$; $nn_{5780}$ - $nn_{5790}$; $nn_{5790}$ - $nn_{5800}$; $nn_{5800}$ - $nn_{5810}$; $nn_{5810}$ - $nn_{5820}$; $nn_{5820}$ - $nn_{5830}$; $nn_{5830}$ - $nn_{5840}$; $nn_{5840}$ - $nn_{5850}$; $nn_{5850}$ - $nn_{5860}$; $nn_{5860}$ - $nn_{5870}$; $nn_{5870}$ - $nn_{5880}$; $nn_{5880}$ - $nn_{5890}$; $nn_{5890}$ - $nn_{5900}$; $nn_{5900}$ - $nn_{5910}$; $nn_{5910}$ - $nn_{5920}$; $nn_{5920}$ - $nn_{5930}$; $nn_{5930}$ - $nn_{5940}$; $nn_{5940}$ - $nn_{5950}$; $nn_{5950}$ - $nn_{5960}$; $nn_{5960}$ - $nn_{5970}$; $nn_{5970}$ - $nn_{5980}$; $nn_{5980}$ - $nn_{5990}$; $nn_{5990}$ - $nn_{6000}$; $nn_{6000}$ - $nn_{6010}$; $nn_{6010}$ - $nn_{6020}$; $nn_{6020}$ - $nn_{6030}$; $nn_{6030}$ - $nn_{6040}$; $nn_{6040}$ - $nn_{6050}$; $nn_{6050}$ - $nn_{6060}$; $nn_{6060}$ - $nn_{6070}$; $nn_{6070}$ - $nn_{6080}$; $nn_{6080}$ - $nn_{6090}$; $nn_{6090}$ - $nn_{6100}$; $nn_{6100}$ - $nn_{6110}$; $nn_{6110}$ - $nn_{6120}$; $nn_{6120}$ - $nn_{6130}$; $nn_{6130}$ - $nn_{6140}$; $nn_{6140}$ - $nn_{6150}$; $nn_{6150}$ - $nn_{6160}$; $nn_{6160}$ - $nn_{6170}$; $nn_{6170}$ - $nn_{6180}$; $nn_{6180}$ - $nn_{6190}$; $nn_{6190}$ - $nn_{6200}$; $nn_{6200}$ - $nn_{6210}$; $nn_{6210}$ - $nn_{6220}$; $nn_{6220}$ - $nn_{6230}$; $nn_{6230}$ - $nn_{6240}$; $nn_{6240}$ - $nn_{6250}$; $nn_{6250}$ - $nn_{6260}$; $nn_{6260}$ - $nn_{6270}$; $nn_{6270}$ - $nn_{6280}$; $nn_{6280}$ - $nn_{6290}$; $nn_{6290}$ - $nn_{6300}$; $nn_{6300}$ - $nn_{6310}$; $nn_{6310}$ - $nn_{6320}$; $nn_{6320}$ - $nn_{6330}$; $nn_{6330}$ - $nn_{6340}$; $nn_{6340}$ - $nn_{6350}$; $nn_{6350}$ - $nn_{6360}$; $nn_{6360}$ - $nn_{6370}$; $nn_{6370}$ - $nn_{6380}$; $nn_{6380}$ - $nn_{6390}$; $nn_{6390}$ - $nn_{6400}$; $nn_{6400}$ - $nn_{6410}$; $nn_{6410}$ - $nn_{6420}$; $nn_{6420}$ - $nn_{6430}$; $nn_{6430}$ - $nn_{6440}$; $nn_{6440}$ - $nn_{6450}$; $nn_{6450}$ - $nn_{6460}$; $nn_{6460}$ - $nn_{6470}$; $nn_{6470}$ - $nn_{6480}$; $nn_{6480}$ - $nn_{6490}$; $nn_{6490}$ - $nn_{6500}$; $nn_{6500}$ - $nn_{6510}$; $nn_{6510}$ - $nn_{6520}$; $nn_{6520}$ - $nn_{6530}$; $nn_{6530}$ - $nn_{6540}$; $nn_{6540}$ - $nn_{6550}$; $nn_{6550}$ - $nn_{6560}$; $nn_{6560}$ - $nn_{6570}$; $nn_{6570}$ - $nn_{6580}$; $nn_{6580}$ - $nn_{6590}$; $nn_{6590}$ - $nn_{6600}$; $nn_{6600}$ - $nn_{6610}$; $nn_{6610}$ - $nn_{6620}$; $nn_{6620}$ - $nn_{6630}$; $nn_{6630}$ - $nn_{6640}$; $nn_{6640}$ - $nn_{6650}$; $nn_{6650}$ - $nn_{6660}$; $nn_{6660}$ - $nn_{6670}$; $nn_{6670}$ - $nn_{6680}$; $nn_{6680}$ - $nn_{6690}$; $nn_{6690}$ - $nn_{6700}$; $nn_{6700}$ - $nn_{6710}$;

$nn_{6710}$ - $nn_{6720}$; $nn_{6720}$ - $nn_{6730}$; $nn_{6730}$ - $nn_{6740}$; $nn_{6740}$ - $nn_{6750}$; $nn_{6750}$ - $nn_{6760}$; $nn_{6760}$ - $nn_{6770}$; $nn_{6770}$ - $nn_{6780}$; $nn_{6780}$ - $nn_{6790}$; $nn_{6790}$ - $nn_{6800}$; $nn_{6800}$ - $nn_{6810}$; $nn_{6810}$ - $nn_{6820}$; $nn_{6820}$ - $nn_{6830}$; $nn_{6830}$ - $nn_{6840}$; $nn_{6840}$ - $nn_{6850}$; $nn_{6850}$ - $nn_{6860}$; $nn_{6860}$ - $nn_{6870}$; $nn_{6870}$ - $nn_{6880}$; $nn_{6880}$ - $nn_{6890}$; $nn_{6890}$ - $nn_{6900}$; $nn_{6900}$ - $nn_{6910}$; $nn_{6910}$ - $nn_{6920}$; $nn_{6920}$ - $nn_{6930}$; $nn_{6930}$ - $nn_{6940}$; $nn_{6940}$ - $nn_{6950}$; $nn_{6950}$ - $nn_{6960}$; $nn_{6960}$ - $nn_{6970}$; $nn_{6970}$ - $nn_{6980}$; $nn_{6980}$ - $nn_{6990}$; $nn_{6990}$ - $nn_{7000}$; $nn_{7000}$ - $nn_{7010}$; $nn_{7010}$ - $nn_{7020}$; $nn_{7020}$ - $nn_{7030}$; $nn_{7030}$ - $nn_{7040}$; $nn_{7040}$ - $nn_{7050}$; $nn_{7050}$ - $nn_{7060}$; $nn_{7060}$ - $nn_{7070}$; $nn_{7070}$ - $nn_{7080}$; $nn_{7080}$ - $nn_{7090}$; $nn_{7090}$ - $nn_{7100}$; $nn_{7100}$ - $nn_{7110}$; $nn_{7110}$ - $nn_{7120}$; $nn_{7120}$ - $nn_{7130}$; $nn_{7130}$ - $nn_{7140}$; $nn_{7140}$ - $nn_{7150}$; $nn_{7150}$ - $nn_{7160}$; $nn_{7160}$ - $nn_{7170}$; $nn_{7170}$ - $nn_{7180}$; $nn_{7180}$ - $nn_{7190}$; $nn_{7190}$ - $nn_{7200}$; $nn_{7200}$ - $nn_{7210}$; $nn_{7210}$ - $nn_{7220}$; $nn_{7220}$ - $nn_{7230}$; $nn_{7230}$ - $nn_{7240}$; $nn_{7240}$ - $nn_{7250}$; $nn_{7250}$ - $nn_{7260}$; $nn_{7260}$ - $nn_{7270}$; $nn_{7270}$ - $nn_{7280}$; $nn_{7280}$ - $nn_{7290}$; $nn_{7290}$ - $nn_{7300}$; $nn_{7300}$ - $nn_{7310}$; $nn_{7310}$ - $nn_{7320}$; $nn_{7320}$ - $nn_{7330}$; $nn_{7330}$ - $nn_{7340}$; $nn_{7340}$ - $nn_{7350}$; $nn_{7350}$ - $nn_{7360}$; $nn_{7360}$ - $nn_{7370}$; $nn_{7370}$ - $nn_{7380}$; $nn_{7380}$ - $nn_{7390}$; $nn_{7390}$ - $nn_{7400}$; $nn_{7400}$ - $nn_{7410}$; $nn_{7410}$ - $nn_{7420}$; $nn_{7420}$ - $nn_{7430}$; $nn_{7430}$ - $nn_{7440}$; $nn_{7440}$ - $nn_{7450}$; $nn_{7450}$ - $nn_{7460}$; $nn_{7460}$ - $nn_{7470}$; $nn_{7470}$ - $nn_{7480}$; $nn_{7480}$ - $nn_{7490}$; $nn_{7490}$ - $nn_{7500}$; $nn_{7500}$ - $nn_{7510}$; $nn_{7510}$ - $nn_{7520}$; $nn_{7520}$ - $nn_{7530}$; $nn_{7530}$ - $nn_{7540}$; $nn_{7540}$ - $nn_{7550}$; $nn_{7550}$ - $nn_{7560}$; $nn_{7560}$ - $nn_{7570}$; $nn_{7570}$ - $nn_{7580}$; $nn_{7580}$ - $nn_{7590}$; $nn_{7590}$ - $nn_{7600}$; $nn_{7600}$ - $nn_{7610}$; $nn_{7610}$ - $nn_{7620}$; $nn_{7620}$ - $nn_{7630}$; $nn_{7630}$ - $nn_{7640}$; $nn_{7640}$ - $nn_{7650}$; $nn_{7650}$ - $nn_{7660}$; $nn_{7660}$ - $nn_{7670}$; $nn_{7670}$ - $nn_{7680}$; $nn_{7680}$ - $nn_{7690}$; $nn_{7690}$ - $nn_{7700}$; $nn_{7700}$ - $nn_{7710}$; $nn_{7710}$ - $nn_{7720}$; $nn_{7720}$ - $nn_{7730}$; $nn_{7730}$ - $nn_{7740}$; $nn_{7740}$ - $nn_{7750}$; $nn_{7750}$ - $nn_{7760}$; $nn_{7760}$ - $nn_{7770}$; $nn_{7770}$ - $nn_{7780}$; $nn_{7780}$ - $nn_{7790}$;

$nn_{7790}$ - $nn_{7800}$; $nn_{7800}$ - $nn_{7810}$; $nn_{7810}$ - $nn_{7820}$; $nn_{7820}$ - $nn_{7830}$; $nn_{7830}$ - $nn_{7840}$; $nn_{7840}$ - $nn_{7850}$; $nn_{7850}$ - $nn_{7860}$; $nn_{7860}$ - $nn_{7870}$; $nn_{7870}$ - $nn_{7880}$; $nn_{7880}$ - $nn_{7890}$; $nn_{7890}$ - $nn_{7900}$; $nn_{7900}$ - $nn_{7910}$; $nn_{7910}$ - $nn_{7920}$; $nn_{7920}$ - $nn_{7930}$; $nn_{7930}$ - $nn_{7940}$; $nn_{7940}$ - $nn_{7950}$; $nn_{7950}$ - $nn_{7960}$; $nn_{7960}$ - $nn_{7970}$; $nn_{7970}$ - $nn_{7980}$; $nn_{7980}$ - $nn_{7990}$; $nn_{7990}$ - $nn_{8000}$; $nn_{8000}$ - $nn_{8010}$; $nn_{8010}$ - $nn_{8020}$; $nn_{8020}$ - $nn_{8030}$; $nn_{8030}$ - $nn_{8040}$; $nn_{8040}$ - $nn_{8050}$; $nn_{8050}$ - $nn_{8060}$; $nn_{8060}$ - $nn_{8070}$; $nn_{8070}$ - $nn_{8080}$; $nn_{8080}$ - $nn_{8090}$; $nn_{8090}$ - $nn_{8100}$; $nn_{8100}$ - $nn_{8110}$; $nn_{8110}$ - $nn_{8120}$; $nn_{8120}$ - $nn_{8130}$; $nn_{8130}$ - $nn_{8140}$; $nn_{8140}$ - $nn_{8150}$; $nn_{8150}$ - $nn_{8160}$; $nn_{8160}$ - $nn_{8170}$; $nn_{8170}$ - $nn_{8180}$; $nn_{8180}$ - $nn_{8190}$; $nn_{8190}$ - $nn_{8200}$; $nn_{8200}$ - $nn_{8210}$; $nn_{8210}$ - $nn_{8220}$; $nn_{8220}$ - $nn_{8230}$; $nn_{8230}$ - $nn_{8240}$; $nn_{8240}$ - $nn_{8250}$; $nn_{8250}$ - $nn_{8260}$; $nn_{8260}$ - $nn_{8270}$; $nn_{8270}$ - $nn_{8280}$; $nn_{8280}$ - $nn_{8290}$; $nn_{8290}$ - $nn_{8300}$; $nn_{8300}$ - $nn_{8310}$; $nn_{8310}$ - $nn_{8320}$; $nn_{8320}$ - $nn_{8330}$; $nn_{8330}$ - $nn_{8340}$; $nn_{8340}$ - $nn_{8350}$; $nn_{8350}$ - $nn_{8360}$; $nn_{8360}$ - $nn_{8370}$; $nn_{8370}$ - $nn_{8380}$; $nn_{8380}$ - $nn_{8390}$; $nn_{8390}$ - $nn_{8400}$; $nn_{8400}$ - $nn_{8410}$; $nn_{8410}$ - $nn_{8420}$; $nn_{8420}$ - $nn_{8430}$; $nn_{8430}$ - $nn_{8440}$; $nn_{8440}$ - $nn_{8450}$; $nn_{8450}$ - $nn_{8460}$; $nn_{8460}$ - $nn_{8470}$; $nn_{8470}$ - $nn_{8480}$; $nn_{8480}$ - $nn_{8490}$; $nn_{8490}$ - $nn_{8500}$; $nn_{8500}$ - $nn_{8510}$; $nn_{8510}$ - $nn_{8520}$; $nn_{8520}$ - $nn_{8530}$; $nn_{8530}$ - $nn_{8540}$; $nn_{8540}$ - $nn_{8550}$; $nn_{8550}$ - $nn_{8560}$; $nn_{8560}$ - $nn_{8570}$; $nn_{8570}$ - $nn_{8580}$; $nn_{8580}$ - $nn_{8590}$; $nn_{8590}$ - $nn_{8600}$; $nn_{8600}$ - $nn_{8610}$; $nn_{8610}$ - $nn_{8620}$; $nn_{8620}$ - $nn_{8630}$; $nn_{8630}$ - $nn_{8640}$; $nn_{8640}$ - $nn_{8650}$; $nn_{8650}$ - $nn_{8660}$; $nn_{8660}$ - $nn_{8670}$; $nn_{8670}$ - $nn_{8680}$; $nn_{8680}$ - $nn_{8690}$; $nn_{8690}$ - $nn_{8700}$; $nn_{8700}$ - $nn_{8710}$; $nn_{8710}$ - $nn_{8720}$; $nn_{8720}$ - $nn_{8730}$; $nn_{8730}$ - $nn_{8740}$; $nn_{8740}$ - $nn_{8750}$; $nn_{8750}$ - $nn_{8760}$; $nn_{8760}$ - $nn_{8770}$; $nn_{8770}$ - $nn_{8780}$; $nn_{8780}$ - $nn_{8790}$; $nn_{8790}$ - $nn_{8800}$; $nn_{8800}$ - $nn_{8810}$; $nn_{8810}$ - $nn_{8820}$; $nn_{8820}$ - $nn_{8830}$; $nn_{8830}$ - $nn_{8840}$; $nn_{8840}$ - $nn_{8850}$; $nn_{8850}$ - $nn_{8860}$; $nn_{8860}$ - $nn_{8870}$;

$nn_{8870}$ - $nn_{8880}$; $nn_{8880}$ - $nn_{8890}$; $nn_{8890}$ - $nn_{8900}$; $nn_{8900}$ - $nn_{8910}$; $nn_{8910}$ - $nn_{8920}$; $nn_{8920}$ - $nn_{8930}$; $nn_{8930}$ - $nn_{8940}$; $nn_{8940}$ - $nn_{8950}$; $nn_{8950}$ - $nn_{8960}$; $nn_{8960}$ - $nn_{8970}$; $nn_{8970}$ - $nn_{8980}$; $nn_{8980}$ - $nn_{8990}$; $nn_{8990}$ - $nn_{9000}$; $nn_{9000}$ - $nn_{9010}$; $nn_{9010}$ - $nn_{9020}$; $nn_{9020}$ - $nn_{9030}$; $nn_{9030}$ - $nn_{9040}$; $nn_{9040}$ - $nn_{9050}$; $nn_{9050}$ - $nn_{9060}$.

The oligomer, however, need not consist only of the sequence which is complementary to the targeted HCV sequence. It may contain in addition, nucleotide sequences or other moieties which are suitable for the purposes for which the oligomers are used. For example, if the oligomers are used as primers for the amplification of HCV sequences via PCR, they may contain sequences which, when in duplex, form restriction enzyme sites which facilitate the cloning of the amplified sequences. For example, also, if the oligomers are to be used as "capture probes" in hybridization assays (described infra), they would contain in addition a binding partner which is coupled to the oligomer containing the nucleotide sequenc which is complementary to the targeted HCV sequence. Other types of moieities or sequences which are useful of which the oligomers may be comprised or coupled to, are those which are known in the art to be suitable for a variety of purposes, including the labeling of nucleotide probes.

The preparation of the oligomers is by means known in the art, including, for example, by methods which include excision, transcription, or chemical synthesis. The target sequences and/or regions of the genome which are selected to which the targeting polynucleotides of the oligomers are complementary depend upon the purpose. For example, if the goal is to screen for the presence of HCV in biological samples (e.g. blood), the preferred oligomers would be used as probes and/or primers, and

would hybridiz to conserved regions of th HCV g n me. Some of the conserved regions of th HCV genome to which the oligom rs may bind are described herein, for example, the r gions which include nucleotide numbers from about the 5-terminus to about 200, or from about 4000 to about 5000, or from about 8000 to about 9040 as shown in Fig. 18, or preferably nucleotides -318 to 174, 4056 to 4448, and 4378 to 4902. Other regions of the genome which are conserved are readily ascertainable by comparison of th nucleotide sequences of various isolates of HCV, including the prototype HCV, HCV1. Methods for conducting comparisons between genotypes to determine conserved and nonconserved regions are known in the art, and examples of these methods are disclosed herein.

In the basic nucleic acid hybridization assay, single-stranded analyte nucleic acid (either DNA or RNA) is hybridized to a nucleic acid probe, and resulting duplexes are detected. The probes for HCV polynucleotides (natural or derived) are a length which allows the detection of unique viral sequences by hybridization. While 6-8 nucleotides may be a workable length, sequences of 10-12 nucleotides are preferred, and about 20 nucleotides or more appears optimal. Preferably, these sequences will derive from regions which lack heterogeneity. These probes can be prepared using routine methods, including automated oligonucleotide synthetic methods. Among useful probes, for example, are those derived from the newly isolated clones disclosed herein, as well as the various oligomers useful in probing cDNA libraries, set forth below. A complement to any unique portion of the HCV genome will be satisfactory. For use as probes, complete complementarity is desirable, though it may be unnecessary as the length of the fragment is increased.

For use of such probes as agents to detect the presence of HCV polynucleotides (for example in screening for contaminated blood), the biological sample to be

analyzed, such as blood or serum, may b treat d, if desir d, to extract the nucl ic acids contain d therein. Th r sulting nucl ic acid from the sample may be subjected to g l lectrophoresis or other size separation techniques; alternatively, the nucleic acid sample may b dot blotted without size separation. In order to form hybrid duplexes with the targeting sequence of the probe, the targeted region of the analyte nucleic acid must be in single stranded form. Where the sequence is naturally present in single stranded form, denaturation will not be required. However, where the sequence is present in double stranded form, the sequence will be denatured. Denaturation can be carried out by various techniques known in the art. Subsequent to denaturation, the analyte nucleic acid and probe are incubated under conditions which promote stable hybrid formation of the target sequence in the probe with the putative targeted sequence in the analyte, and the resulting duplexes containing the probe(s) are detected.

Detection of the resulting duplex, if any, is usually accomplished by the use of labeled probes; alternatively, the probe may be unlabeled, but may be detectable by specific binding with a ligand which is labeled, either directly or indirectly. Suitable labels, and methods for labeling probes and ligands are known in the art, and include, for example, radioactive labels which may be incorporated by known methods (e.g., nick translation or kinasing), biotin, fluorescent groups, chemiluminescent groups (e.g., dioxetanes, particularly triggered dioxetanes), enzymes, antibodies, and the lik .

The region of the probes which are used to bind to the analyte can be made completely complementary to the HCV genome. Therefore, usually high stringency conditions are desirable in order to prevent false positives. However, conditions of high stringency should only be used if the probes are complementary to regions of the viral

genom which lack h terogeneity. The stringency of hybridization is determined by a number of factors during hybridization and during the washing procedure, including temperature, ionic strength, length of time, and concentration of formamide. These factors are outlined in, for example, Maniatis, T. (1982).

Variations of this basic scheme which are known in the art, including those which facilitate separation of the duplexes to be detected from extraneous materials and/or which amplify the signal from the labeled moiety, may also be used. A number of these variations are reviewed in, for example: Matthews and Kricka (1988), Analytical Biochemistry 169:1; Landegren et al. (1988), Science 242:229; and Mittlin (1989), Clinical chem. 35:1819. These and the following publications describing assay formats are hereby incorporated by reference herein. Probes suitable for detecting HCV in these assays are comprised of sequences which hybridize with target HCV polynucleotide sequences to form duplexes with the analyte strand, wherein the duplexes are of sufficient stability for detection in the specified assay system.

A suitable variation is, for example, one which is described in U.S. Patent No. 4,868,105, issued Sept. 9, 1989, and in E.P.O. Publication No. 225,807 (published June 16, 1987). These publications describe a solution phase nucleic acid hybridization assay in which the analyte nucleic acid is hybridized to a labeling probe set and to a capturing probe set. The probe-analyte complex is coupled by hybridization with a solid-supported capture probe that is complementary to the capture probe set. This permits the analyte nucleic acid to be removed from solution as a solid phase complex. Having the analyte in the form of a solid phase complex facilitates subsequent separation steps in the assay. The labeling probe set is complementary to a labeled probe that is bound through hybridization to the solid phase/analyte complex.

Gen rally, it is xp cted that the HCV genome sequences will b pres nt in serum of infect d individuals at relativ ly low lev ls, i. ., at approximately $10^2$-$10^3$ chimp infectious doses (CID) per ml. This level may require that amplification techniques be used in hybridization assays. Such techniques are known in th art. For example, the Enzo Biochemical Corporation "Bio-Bridge" system uses terminal deoxynucleotide transferase to add unmodified 3'-poly-dT-tails to a DNA probe. The poly dT-tailed probe is hybridized to the target nucleotide sequence, and then to a biotin-modified poly-A. PCT Publication 84/03520 and EP Publication No. 124221 describe a DNA hybridization assay in which: (1) analyte is annealed to a single-stranded DNA probe that is complementary to an enzyme-labeled oligonucleotide; and (2) the resulting tailed duplex is hybridized to an enzyme-labeled oligonucleotide. EPA 204510 describes a DNA hybridization assay in which analyte DNA is contacted with a probe that has a tail, such as a poly-dT tail, an amplifier strand that has a sequence that hybridizes to the tail of the probe, such as a poly-A sequence, and which is capable of binding a plurality of labeled strands. A type of hybridization assay which is described in E.P.O. Publication No. 317,077 (published May 24, 1989), which should detect sequences at the level of approximately $10^6$/ml, utilizes nucleic acid multimers which bind to single-stranded analyte nucleic acid, and which also bind to a multiplicity of single-stranded labeled oligonucleotides. A particularly desirable technique may involve amplification of the target HCV sequences in sera approximately 10,000 fold (i.e., to approximately $10^6$ sequences/ml), as part of the hybridization system. The amplification may be accomplished, for example, by the polymerase chain reactions (PCR) technique described by Saiki et al. (1986), by Mullis, U.S. Patent No. 4,683,195, and by Mullis et al. U.S. Patent No. 4,683,202. Amplification may be prior to,

or pr f rably subs quent to purification of the HCV target s quence. For example, amplification may be utilized in conjunction with th assay m thods described in U.S. Pat- nt No. 4,868,105, or if even further amplification is desired, in conjunction with the hybridization system described in E.P.O. Publication No. 317,077.

Preferred methods for detecting HCV sequences in an analyte polynucleotide strand are based upon the hybridization detection methods described in U.S. Patent No. 4,868,105 and in E.P.O. Publication No. 317,077. These methods are solution-phase sandwich hybridization assays which utilize both capture and label probes which hybridize to target sequences in an analyte nucleic acid. In the use of these assays to screen biological samples for HCV, the probes used would bind to conserved regions of the HCV genome. The capture and label probes may be interspersed in their binding to the target sequence. Alternatively, in a preferred mode the capture and label probes are in sets, and the probes of one set do not intersperse with the probes of another set. In the latt r mode, preferably the set(s) of multiple capture probes hybridize to the most conserved regions of the genome, while the set(s) of multiple label probes may hybridize to regions which exhibit small amounts of divergence. For example, using the prototype HCV1 cDNA sequence shown in Fig. 18, probes could be used which hybridize to sequences in the region of nucleotides from about -318 to about 174, and/or nucleotides in the region of about 4378 to about 4902, and/or nucleotides in the region of from about 4056 to about 4448. The preferred probes would hybridize to sequences in the 5′-region of the HCV genome, since, as shown infra., this region appears to be highly conserved. Thus, preferred probes may hybridize to, for example, nucleotides from about -318 to about 174 as shown in Fig. 18. Probes could be used which hybridize to either the positive strand in conserved regions, and/or its comple-

ment, d pending upon th purpos , for example, to d tect viral genomic sequences, or to det ct HCV cDNA sequences resulting from PCR amplification, or to d t ct replicative intermediates to the positive HCV RNA strand.

Detection of HCV RNA and Polynucleotides Derived Therefrom Using an HCV/cPCR Method

A particularly useful method for detecting HCV RNA or polynucleotides derived from HCV RNA is the HCV/cPCR method, which is a subject of the herein application, and which utilizes the polymerase chain reaction technique (PCR) which is described by Saiki et al. (1986), by Mullis in U.S. Pat. No. 4,683,195, and by Mullis et al. in U.S. Patent No. 4,683,202. The HCV/cPCR method utilizes primers and probes derived from the information provided herein concerning the nature of the HCV genome.

Generally, in the PCR technique, short oligonucleotide primers are prepared which match opposite ends of a desired sequence. The sequence between the primers need not be known. A sample of polynucleotide is extracted and denatured, preferably by heat, and hybridized with oligonucleotide primers which are present in molar excess. Polymerization is catalyzed by a templat - and primer-dependent polymerase in the presence of deoxynucleotide triphosphates or nucleotide analogs (dNTPs). This results in two "long products" which contain the respective primers at their 5'-termini, covalently linked to the newly synthesized complements of the original strands. The replicated DNA is again denatured, hybridized with oligonucleotide primers, return d to polymerizing conditions, and a second cycle of replication is initiated. The second cycle provides the two original strands, the two long products from cycle 1, and two "short products" replicated from the long products. The short products contain sequences (sense or antisense) derived from the target sequence, flanked at the 5'- and

3'-t rmini with primer sequenc s. On each additional cycle, th number of short products is replicated exponentially. Thus, this process causes the amplification of a specific targ t sequ nce.

In the method, a sample is provided which is suspected of containing HCV RNA, or a fragment thereof. The sample is usually taken from an individual suspected of having NANBH; however, other sources of the sample are included, e.g., conditioned medium or cells from in vitro systems in which the virus has been replicated. The sample, however, must contain the target nucleic acid sequence(s).

The sample is then subjected to conditions which allow reverse transcription of HCV RNA into HCV cDNA. Conditions for reverse transcribing RNA are known to those of skill in the art, and are described in, for example, Maniatis et al. (1982), and in Methods in Enzymology. A preferred method of reverse transcription utilizes reverse transcriptase from a variety of sources, including recombinant molecules, and isolated from, for example, a retrovirus, preferably from avian myeloblastosis virus (AMV), and suitable conditions for the transcription. The HCV cDNA product of reverse transcription is in a RNA:DNA hybrid, which results from the first round of reverse transcription; subsequently, DNA:DNA hybrids result from two or more rounds of transcription.

The HCV cDNA resulting from reverse transcription is then subjected to PCR to amplify the target sequence. In order to accomplish this, the HCV cDNA is denatured, and the separated strands are hybridized with primers which flank the target sequence.

Strand separation may be accomplished by any suitable denaturing method, including physical, chemical, or enzymatic means, which are known to those of skill in the art. A preferred method, which is physical, involves heating the nucleic acid until it is completely (>99%)

denatur d. Typical heat d naturation involves temperatur s ranging from about 80°C to about 105°C, for times ranging from about 1 to 10 minutes.

After hybridization of the HCV cDNA with the primers, the target HCV sequences are replicated by a polymerizing means which utilizes a primer oligonucleotide to initiate the synthesis of the replicate chain. The primers are selected so that they are complementary to sequences of the HCV genome. Oligomeric primers which are complementary to regions of the sense and antisense strands of HCV cDNA can be designed from the HCV cDNA sequences from the composite cDNA sequence provided in Fig. 18.

The primers are selected so that their relative positions along a duplex sequence are such that an extension product synthesized from one primer, when it is separated from its template (complement), serves as a template for the extension of the other primer to yield a replicate chain of defined length.

The primer is preferably single stranded for maximum efficiency in amplification, but may alternatively be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare extension products. Preferably, the primer is an oligodeoxyribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the agent for polymerization. The exact lengths of the primers will depend on many factors, including temperature and source of the primer and use of the method. For example, depending on the complexity of the target sequence, the oligonucleotide primer typically contains about 15-45 nucleotides, although it may contain more or fewer nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with the template.

Th primers used h r in are s lected to be "substantially" complementary to the different strands of each specific sequence to be amplified. Therefore, the primers ne d not reflect the exact sequ nc of th template, but must be sufficiently complementary to selectively hybridize with their respective strands. For example, a non-complementary nucleotide fragment may b attached to the 5'-end of the primer, with the remainder of the primer sequence being complementary to the strand. Alternatively, non-complementary bases or longer sequences can be interspersed into the primer, provided that the primer has sufficient complementarity with the sequence of one of the strands to be amplified to hybridize therewith, and to thereby form a duplex structure which can be extended by the polymerizing means. The non-complementary nucleotide sequences of the primers may include restriction enzyme sites. Appending a restriction enzyme site to the end(s) of the target sequence would be particularly helpful for cloning of the target sequence.

It will be understood that "primer", as used herein, may refer to more than one primer, particularly in the case where there is some ambiguity in the information regarding the terminal sequence(s) of the target region to be amplified. Hence, a "primer" includes a collection of primer oligonucleotides containing sequences representing the possible variations in the sequence or includes nucleotides which allow a typical basepairing. One of the primer oligonucleotides in this collection will be homologous with the end of the target sequence. A specific case is shown in the Examples, where oligomer sets of 44-mers and 45-mers were utilized to prime the amplification of a potentially variant region of the HCV genome.

It is anticipated that there will be a variety of strains or isolates of HCV with sequences which deviate from HCV1, the prototype strain. Therefore, in order to

d t ct variant strains it is preferabl to construct primers which hybridize to cons rv d r gions of the HCV g nom . Th cons rved regions may b determin d by comparing the nucleotid or amino acid sequences of several HCV strains/isolates. There appear to be at least three regions of conserved amino acid in the HCV genom , described supra., from which primers may be derived. These regions are believed to be. The primers described infra., in the Examples, are derived from what are believed to be conserved regions of HCV, based upon sequence homology to that of the Flaviviruses.

The oligonucleotide primers may be prepared by any suitable method. Methods for preparing oligonucleotides of specific sequence are known in the art, and include, for example, cloning and restriction of appropriate sequences, and direct chemical synthesis. Chemical synthesis methods may include, for example, th phosphotriester method described by Narang et al. (1979), the phosphodiester method disclosed by Brown et al. (1979), the diethylphosphoramidate method disclosed in Beaucage et al. (1981), and the solid support method in U.S. Patent No. 4,458,066.

The primers may be labeled, if desired, by incorporating means detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means.

Template-dependent extension of the oligonucleotide primer(s) is catalyzed by a polymerizing agent in the presence of adequate amounts of the four deoxyribonucleotide triphosphates (dATP, dGTP, dCTP and dTTP) or analogs, in a reaction medium which is comprised of the appropriate salts, metal cations, and pH buffering system. Suitable polymerizing agents are enzymes known to catalyze primer- and template-dependent DNA synthesis. Known DNA polymerases include, for example, E. coli DNA polymerase I or its Klenow fragment, $T_4$ DNA polymerase, and Taq DNA polymerase. The reaction conditions for

catalyzing DNA synth sis with th s DNA polym rases are kn wn in th art.

Th products of the synthesis are dupl x mol cul s consisting of the template strands and the primer extension strands, which include the target sequence. These products, in turn, serve as template for another round of replication. In the second round of replication, the primer extension strand of the first cycle is annealed with its complementary primer; synth sis yields a "short" product which is bounded on both the 5'- and the 3'-ends by primer sequences or their complements. Repeated cycles of denaturation, primer annealing, and extension result in the exponential accumulation of the target region defined by the primers. Sufficient cycles are run to achieve the desired amount of polynucleotide containing the target region of nucleic acid. The desir d amount may vary, and is determined by the function which the product polynucleotide is to serve.

The PCR method can be performed in a number of temporal sequences. For example, it can be performed step-wise, where after each step new reagents are added, or in a fashion where all of the reagents are added simultaneously, or in a partial step-wise fashion, where fresh reagents are added after a given number of steps.

In a preferred method, the PCR reaction is carried out as an automated process which utilizes a thermostable enzyme. In this process the reaction mixture is cycled through a denaturing region, a primer annealing region, and a reaction region. A machine may be employed which is specifically adapted for use with a thermostable enzyme, which utilizes temperature cycling without a liquid handling system, since the enzyme need not be added at every cycle. This type of machine is commercially available from Perkin Elmer Cetus Corp.

After amplification by PCR, the target polynucleotides are detected by hybridization with a probe

polynucl otid which forms a stabl hybrid with that of th targ t s quenc und r stringent to moderat ly stringent hybridization and wash conditions. If it is expected that th prob s will b complet ly complementary (i.e., about 99% or greater) to the target sequence, stringent conditions will be used. If some mismatching is expected, for example if variant strains are expected with the result that the probe will not be completely complementary, the stringency of hybridization may be lessened. However, conditions are chosen which rule out nonspecific/adventitious binding. Conditions which affect hybridization, and which select against nonspecific binding are known in the art, and are described in, for example, Maniatis et al. (1982). Generally, lower salt concentration and higher temperature increase the stringency of binding. For example, it is usually considered that stringent conditions are incubation in solutions which contain approximately 0.1 X SSC, 0.1% SDS, at about 65°C incubation/wash temperature, and moderately stringent conditions are incubation in solutions which contain approximately 1-2 X SSC, 0.1% SDS and about 50°-65°C incubation/wash temperature. Low stringency conditions are 2 X SSC and about 30°-50°C.

Probes for HCV target sequences may be deriv d from the HCV cDNA sequence shown in Fig. 18, or from new HCV isolates. The HCV probes may be of any suitable length which span the target region, but which exclude the primers, and which allow specific hybridization to the target region. If there is to be complete complementarity, i.e., if the strain contains a sequence identical to that of the probe, since the duplex will be relatively stable under even stringent conditions, the probes may be short, i.e., in the range of about 10-30 base pairs. If some degree of mismatch is expected with the probe, i.e., if it is suspected that the probe will hybridize to a variant region, the probe may be of greater

l ngth, since length seems to counterbalance some of the effect of th mismatch(es). An exampl of this is found in th Exampl s, where the probe was designed to bind to potential variants of HCV1. In this case, th primers were designed to bind to HCV cDNA derived from a hypothetical conserved region of the HCV genome, and the target region was one which potentially contained variations (based upon the Flavivirus model). The probe used to detect the HCV target sequences contained approximately 268 base pairs.

The probe nucleic acid having a sequence complementary to the target sequence may be synthesized using similar techniques described supra. for the synthesis of primer sequences. If desired, the probe may be labeled. Appropriate labels are described supra.

In some cases, it may be desirable to determine the length of the PCR product detected by the probe. This may be particularly true if it is suspected that variant HCV strains may contain deletions within the target region, or if one wishes to confirm the length of the PCR product. In such cases it is preferable to subject th products to size analysis as well as hybridization with the probe. Methods for determining the size of nucleic acids are known in the art, and include, for example, g l electrophoresis, sedimentation in gradients, and gel exclusion chromatography.

The presence of the target sequence in a biological sample is detected by determining whether a hybrid has been formed between the HCV polynucleotide probe and the nucleic acid subjected to the PCR amplification technique. Methods to detect hybrids formed between a probe and a nucleic acid sequence are known in the art. For example, for convenience, an unlabeled sample may be transferred to a solid matrix to which it binds, and the bound sample subjected to conditions which allow specific hybridization with a labeled probe; the solid matrix is

than xamin d for the presenc of the labeled prob . Alternativ ly, if the sample is label d, th unlabeled probe is bound to the matrix, and after the exposure to th appropriate hybridization conditions, the matrix is examined for the presence of label. Other suitable hybridization assays are described supra.

Determination of Variant HCV Sequences Using PCR

In order to identify variant HCV strains, and thereby to design probes for those variants, the above described HCV/cPCR method is utilized to amplify variant regions of the HCV genome, so that the nucleotide sequences of these variant target regions can be determined. Generally, variant types of HCV might be expected to occur in different geographic locations than that in which the HCV1 strain is predominant, for example, Japan, Africa, etc.; or in different vertebrate species which are also infected with the virus. Variant HCV may also arise during passage in tissue culture systems, or be the result of spontaneous or induced mutations.

In order to amplify the variant target region, primers are designed to flank the suspect region, and preferably are complementary to conserved regions. Primers to two regions of HCV which are probably conserved, based upon the Flavivirus model, are described in the Examples. These primers and probes may be designed utilizing the sequence information for the HCV1 strain provided in Fig. 18.

Analysis of the nucleotide sequence of the target region(s) may be by direct analysis of the PCR amplified products. A process for direct sequence analysis of PCR amplified products is described in Saiki et al. (1988).

Alternatively, the amplified target sequence(s) may be cloned prior to sequence analysis. A method for the direct cloning and sequence analysis of enzymatically

amplifi d g nomic segments has been d scribed by Scharf (1986). In the method, the primers used in the PCR technique are modified near their 5'-ends to produce convenient restriction sit s for cloning directly into, for example, an M13 sequencing vector. After amplification, the PCR products are cleaved with the appropriate restriction enzymes. The restriction fragments are ligated into the M13 vector, and transformed into, for example, a JM 103 host, plated out, and the resulting plaques are screened by hybridization with a labeled oligonucleotide probe. Other methods for cloning and sequence analysis are known in the art.

Universal Primers for Flaviviruses and for HCV

Studies of the nature of the genome of the HCV, utilizing probes derived from the HCV cDNA, as well as sequence information contained within the HCV cDNA, are suggestive that HCV is a Flavi-like virus. These studies are described in E.P.O. publication No. 318,216 owned by the herein assignee, and which is incorporated herein in its entirety. A comparison of the HCV cDNA sequence derived from the HCV cDNA clones with known sequences of a number of Flaviviruses show that HCV contains sequences which are homologous to conserved sequences in the Flaviviruses. These conserved sequences may allow the creation of primers which may be universal in their application for amplification of target regions of Flaviviruses, and for HCV. These sequences are the 16-mer or smaller sequences from the 3'-termini of the primers described in the Examples. Identification of the species is then accomplished utilizing a probe specific for the species. The genomes of a number of Flaviviruses are known in the art, and include, for example, Japanese Encephalitis Virus (Sumiyoshi et al. (1987)), Yellow Fever Virus (Rice et al. (1985)), Dengue Type 2 Virus (Hahn et al. (1988)), Dengue Type 4 Virus (Mackow (1987)), and West

Nil Virus (Castle t al. (1986)). Identificati n of HCV RNA is accomplished utilizing a prob specific for HCV, th sequenc of which can be determined the HCV cDNA s qu nces provid d h r in.

Alternatively, utilization of sets of probe(s) designed to account for codon degeneracy and therefore contain common sequences to the Flaviviruses and to HCV, as determined by a comparison of HCV amino acid sequences with the known sequences of the Flaviviruses, allows a general detection system for these viruses.

Construction of Desired DNA Sequences

Synthetic oligonucleotides may be prepared using an automated oligonucleotide synthesizer as described by Warner (1984). If desired the synthetic strands may be labeled with $^{32}P$ by treatment with polynucleotide kinase in the presence of $^{32}P$-ATP, using standard conditions for the reaction.

DNA sequences, including those isolated from cDNA libraries, may be modified by known techniques, including, for example site directed mutagenesis, as described by Zoller (1982). Briefly, the DNA to be modified is packaged into phage as a single stranded sequence, and converted to a double stranded DNA with DNA polymerase using, as a primer, a synthetic oligonucleotide complementary to the portion of the DNA to be modified, and having the desired modification included in its own sequence. The resulting double stranded DNA is transformed into a phage supporting host bacterium. Cultures of the transformed bacteria, which contain replications of each strand of the phage, are plated in agar to obtain plaques. Theoretically, 50% of the new plaques contain phage having the mutated sequence, and the remaining 50% have the original sequence. Replicates of the plaques are hybridized to labeled synthetic probe at temperatures and conditions which permit hybridization

with the correct strand, but not with the unmodified sequence. The sequ nc s which have been identified by hybridization are recovered and cloned.

Kits for Screening for HCV Derived Polynucleotides

Oligomers which are probes and/or primers for amplification and/or screening of samples for HCV can be packaged into kits. Kits for screening for HCV sequences include the oligomeric probe DNAs. Kits for amplification of HCV sequences may include the oligomeric primers used in the amplification. The kits usually contain the probes or primers in a premeasured or predetermined amount, as well as other suitably packaged reagents and materials, in separate suitable containers, needed for the particular hybridization and/or amplification protocol(s). For example, the kit may contain standards, buffers, supports, enzymes, substrates, label probes, binding partners, and/or instructions for conducting the test.

## Examples

Described below are examples of the present invention which are provided only for illustrative purposes, and not to limit the scope of the present invention.

Isolation and Sequence of Overlapping HCV cDNA Clones 13i, 26j, CA59a, CA84a, CA156e and CA167b

The clones 13i, 26j, CA59a, CA84a, CA156e and CA167b were isolated from the lambda-gt11 library which contains HCV cDNA (ATCC No. 40394), the preparation of which is described in E.P.O. Publication No. 318,216 (published 31 May 1989), and WO 89/04669 (published 1 June 1989). Screening of the library was with the probes described infra., using the method described in Huynh (1985). The frequencies with which positive clones appeared with the respective probes was about 1 in 50,000.

The isolation of clone 13i was accomplished using a synthetic probe deriv d from the sequence of clone 12f. Th sequenc of the probe was:

5' GAA CGT TGC GAT CTG GAA GAC AGG GAC AGG 3'.

The isolation of clone 26j was accomplished using a probe derived from the 5'-region of clone K9-1. The sequence of the probe was:

5' TAT CAG TTA TGC CAA CGG AAG CGG CCC CGA 3'.

The isolation procedures for clone 12f and for clone k9-1 (also called K9-1) are described in E.P.O. Publication No. 318,216, and their sequences are shown in Figs. 1 and 2, respectively. The HCV cDNA sequences of clones 13i and 26j, are shown in Figs. 4 and 5, respectively. Also shown are the amino acids encoded therein, as well as the overlap of clone 13i with clone 12f, and the overlap of clone 26j with clone 13i. The sequences for these clones confirmed the sequence of clone K9-1. Clone K9-1 had been isolated from a different HCV cDNA library (See E.P.O. Publication No. 218,316).

Clone CA59a was isolated utilizing a probe based upon the sequence of the 5'-region of clone 26j. The sequence of this probe was:

5' CTG GTT AGC AGG GCT TTT CTA TCA CCA CAA 3'.

.A probe derived from the sequence of clone CA59a was used to isolate clone CA84a. The sequence of the probe used for this isolation was:

5' AAG GTC CTG GTA GTG CTG CTG CTA TTT GCC 3'.

Clone CA156e was isolated using a probe derived from th sequenc of clone CA84a. The sequence of the prob was:

5' ACT GGA CGA CGC AAG GTT GCA ATT GCT CTA 3'.

Clone CA167b was isolated using a probe derived from the sequence of clone CA 156e. The sequence of the probe was:

5' TTC GAC GTC ACA TCG ATC TGC TTG TCG GGA 3'.

The nucleotide sequences of the HCV cDNAs in clones CA59a, CA84a, CA156e, and CA167b, are shown Figs. 6, 7, 8, and 9, respectively. The amino acids encoded therein, as well as the overlap with the sequences of relevant clones, are also shown in the figures.

Creation of "pi" HCV cDNA Library

A library of HCV cDNA, the "pi" library, was constructed from the same batch of infectious chimpanzee plasma used to construct the lambda-gt11 HCV cDNA library (ATCC No. 40394) described in E.P.O. Publication No. 318,216, and utilizing essentially the same techniques. However, construction of the pi library utilized a primer-extension method, in which the primer for reverse transcriptase was based on the sequence of clone CA59a. The sequence of the primer was:

5' GGT GAC GTG GGT TTC 3'.

Isolation and Sequence of Clone pi14a

Screening of the "pi" HCV cDNA library described supra., with the probe used to isolate clone CA167b (See supra.) yielded clone pi14a. The clone contains about 800 base pairs of cDNA which overlaps clones CA167b, CA156e,

CA84a and CA59a, which were isolated from th lambda gt-11 HCV cDNA library (ATCC No. 40394). In addition, pi14a also contains about 250 bas pairs of DNA which ar upstream of th HCV cDNA in clone CA167b.

Isolation and Sequence of Clones CA216a, CA290a and ag30a

Based on the sequence of clone CA167b a synthetic probe was made having the following sequence:

5′ GGC TTT ACC ACG TCA CCA ATG ATT GCC CTA 3′

The above probe was used to screen the , which yielded clone CA216a, whose HCV sequences are shown in Fig. 10.

Another probe was made based on the sequence of clone CA216a having the following sequence:

5′ TTT GGG TAA GGT CAT CGA TAC CCT TAC GTG 3′

Screening the lambda-gt11 library (ATCC No. 40394) with this probe yielded clone CA290a, the HCV sequences ther in being shown in Fig. 11.

In a parallel approach, a primer-extension cDNA library was made using nucleic acid extracted from the same infectious plasma used in the original lambda-gt11 cDNA library described above. The primer used was based on the sequence of clones CA216a and CA290a:

5′ GAA GCC GCA CGT AAG 3′

The cDNA library was made using methods similar to those described previously for libraries used in the isolation of clones pi14a and k9-1. The probe used to screen this library was based on the sequence of clone CA290a:

5′ CCG GCG TAG GTC GCG CAA TTT GGG TAA 3′

Clon ag30a was isolated fr m the new library with the above probe, and contained about 670 basepairs of HCV s qu nc . See Fig. 12. Part of this sequence ov rlaps the HCV sequence of clones CA216a and CA290a. About 300 base-pairs of the ag30a sequence, however, is upstream of the sequence from clone CA290a. The non-overlapping sequence shows a start codon (*) and stop codons that may indicate the start of the HCV ORF. Also indicated in Fig. 12 are putative small encoded peptides (#) which may play a role in regulating translation, as well as the putative first amino acid of the putative polypeptide (/), and downstream amino acids encoded therein.

Isolation and Sequence of Clone CA205a

Clone CA205a was isolated from the original lambda gt-11 library (ATCC No. 40394), using a synthetic probe derived from the HCV sequence in clone CA290a (Fig. 11). The sequence of the probe was:

5' TCA GAT CGT TGG TGG AGT TTA CTT GTT GCC 3'.

The sequence of the HCV cDNA in CA205a, shown in Fig. 13, overlaps with the cDNA sequences in both clones ag30a and CA290a. The overlap of the sequence with that of CA290a is shown by the dotted line above the sequence (the figure also shows the putative amino acids encoded in this fragment).

As observed from the HCV cDNA sequences in clones CA205a and ag30a, the putative HCV polyprotein appears to begin at the ATG start codon; the HCV sequences in both clones contain an in-frame, contiguous double stop codon (TGATAG) forty two nucleotides upstream from this ATG. The HCV ORF appears to begin after these stop codons, and to extend for at least 8907 nucleotides (See the composite HCV cDNA shown in Fig. 18).

Isolation and Sequenc of Clone 18g

Based on th sequenc of clone ag30a (See Fig. 12) and of an overlapping clone from the original lambda gt-11 library (ATCC No. 40394), CA230a, a synthetic probe was made having the following sequence:

5′ CCA TAG TGG TCT GCG GAA CCG GTG AGT ACA 3′.

Screening of the original lambda-gt11 HCV cDNA library with the probe yielded clone 18g, the HCV cDNA sequence of which is shown in Fig. 14. Also shown in the figure are the overlap with clone ag30a, and putative polypeptides encoded within the HCV cDNA.

The cDNA in clone 18g (C18g or 18g) overlaps that in clones ag30a and CA205a, described supra. The sequence of C18g also contains the double stop codon region observed in clone ag30a. The polynucleotide region upstream of these stop codons presumably represents part of the 5′-region of the HCV genome, which may contain short ORFs, and which can be confirmed by direct sequencing of the purified HCV genome. These putative small encoded peptides may play a regulatory role in translation. The region of the HCV genome upstream of that represented by C18g can be isolated for sequence analysis using essentially the technique described in E.P.O. Publication No. 318,216 for isolating cDNA sequences upstream of the HCV cDNA sequence in clone 12f. Essentially, small synthetic oligonucleotide primers of reverse transcriptase, which are based upon the sequence of C18g, are synthesized and used to bind to the corresponding sequence in HCV genomic RNA. The primer sequences are proximal to the known 5′-terminal of C18g, but sufficiently downstream to allow the design of probe sequences upstream of the primer sequences. Known standard methods of priming and cloning ar eused. The

r sulting cDNA libraries are scr en d with sequences upstream f th priming sit s (as deduced from th elucidated sequence of C18g). The HCV genomic RNA is obtain d from either plasma or liver samples from individuals with NANBH. Since HCV appears to be a Flavi-like virus, the 5'-terminus of the genome may be modified with a "cap" structure. It is known that Flavivirus genomes contain 5'-terminal "cap" structures. (Yellow Fever virus, Rice et al. (1988); Dengue virus, Hahn et al (1988); Japanese Encephalitis Virus (1987)).

Isolation and Sequence of Clones from the beta-HCV cDNA library

Clones containing cDNA representative of the 3'-terminal region of the HCV genome were isolated from a cDNA library constructed from the original infectious chimpanzee plasma pool which was used for the creation of the HCV cDNA lambda-gt11 library (ATCC No. 40394), described in E.P.O. Publication No. 318,216. In order to create the DNA library, RNA extracted from the plasma was "tailed" with poly rA using poly (rA) polymerase, and cDNA was synthesized using oligo$(dT)_{12-18}$ as a primer for reverse transcriptase. The resulting RNA:cDNA hybrid was digested with RNAase H, and converted to double stranded HCV cDNA. The resulting HCV cDNA was cloned into lambda-gt10, using essentially the technique described in Huynh (1985), yielding the beta (or b) HCV cDNA library. The procedures used were as follows.

An aliquot (12ml) of the plasma was treated with proteinase K, and extracted with an equal volume of phenol saturated with 0.05M Tris-Cl, pH 7.5, 0.05% (v/v) beta-mercaptoethanol, 0.1% (w/v) hydroxyquinolone, 1 mM EDTA. The resulting aqueous phase was re-extracted with the phenol mixture, followed by 3 extractions with a 1:1 mixture containing phenol and chloroform:isoamyl alcohol (24:1), followed by 2 extractions with a mixture of

chl roform and isoamyl alcohol (1:1). Subs quent to adjustm nt of th aqueous phase to 200 mM with respect to NaCl, nucl ic acids in th aqueous phase wer pr cipitated ov rnight at $-20^{o}C$, with 2.5 volum s of cold absolute ethanol. The precipitates were collected by centrifugation at 10,000 RPM for 40 min., washed with 70% ethanol containing 20 mM NaCl, and with 100% cold ethanol, dried for 5 min. in a dessicator, and dissolved in water.

The isolated nucleic acids from the infectious chimpanzee plasma pool were tailed with poly rA utilizing poly-A polymerase in the presence of human placenta ribonuclease inhibitor (HPRI) (purchased from Amersham Corp.), utilizing MS2 RNA as carrier. Isolated nucleic acids equivalent to that in 2 ml of plasma were incubated in a solution containing TMN (50 mM Tris HCl, pH 7.9, 10 mM $MgCl_2$, 250 mM NaCl, 2.5 mM $MnCl_2$, 2 mM dithiothreitol (DTT)), 40 micromolar alpha-[$^{32}P$] ATP, 20 units HPRI (Amersham Corp.), and about 9 to 10 units of RNase fre poly-A polymerase (BRL). Incubation was for 10 min. at $37^{o}C$, and the reactions were stopped with EDTA (final concentration about 250 mM). The solution was extracted with an equal volume of phenol-chloroform, and with an equal volume of chloroform, and nucleic acids were precipitated overnight at $-20^{o}C$ with 2.5 volumes of ethanol in the presence of 200 mM NaCl.

Isolation of Clone b5a

The beta HCV cDNA library was screened by hybridization using a synthetic probe, which had a sequence based upon the HCV cDNA sequence in clone 15e. The isolation of clone 15e is described in E.P.O. Publication No. 318,216, and its sequence is shown in Fig. 3. The sequence of the synthetic probe was:

5' ATT GCG AGA TCT ACG GGG CCT GCT ACT CCA 3'.

Screening of th library yielded clone beta-5a (b5a), which contains an HCV cDNA region of approximat ly 1000 base pairs. The 5'-region of this cDNA overlaps clones 35f, 19g, 26g, and 15e (these clon s are described supra). The region between the 3'-terminal poly-A sequence and the 3'-sequence which overlaps clone 15e, contains approximately 200 base pairs. This clone allows the identification of a region of the 3'-terminal sequence the HCV genome.

The sequence of b5a is contained within the sequence of the HCV cDNA in clone 16jh (described infra). Moreover, the sequence is also present in CC34a, isolat d from the original lambda-gt11 library (ATCC No. 40394). (The original lambda-gt11 library is referred to herein as the "C" library).

Isolation and Sequence of Clones Generated by PCR Amplification of the 3'-Region of the HCV Genome

Multiple cDNA clones have been generated which contain nucleotide sequences derived from the 3'-region of the HCV genome. This was accomplished by amplifying a targeted region of the genome by a polymerase chain reaction technique described in Saiki et al. (1986), and in Saiki et al. (1988), which was modified as described below. The HCV RNA which was amplified was obtained from the original infectious chimpanzee plasma pool which was used for the creation of the HCV cDNA lambda-gt11 library (ATCC No. 40394) described in E.P.O. Publication No. 318,216. Isolation of the HCV RNA was as described supra. The isolated RNA was tailed at the 3'-end with ATP by *E. coli* poly-A polymerase as described in Sippel (1973), except that the nucleic acids isolated from chimp serum were substituted for the nucleic acid substrate. The tailed RNA was then reverse transcribed into cDNA by reverse transcriptase, using an oligo dT-primer adapter,

essentially as described by Han (1987), except that the components and sequ nce of the primer-adapter w re:

| Stuffer | NotI | SP6 Promoter | Primer |
|---|---|---|---|
| AATTC | GCGGCCGC | CATACGATTTAGGTGACACTATAGAA | $T_{15}$ |

The resultant cDNA was subjected to amplification by PCR using two primers:

| Primer | Sequence |
|---|---|
| JH32 (30mer) | ATAGCGGCCGCCCTCGATTGCGAGATCTAC |
| JH11 (20mer) | AATTCGGGCGGCCGCCATACGA |

The JH32 primer contained 20 nucleotide sequences hybridizable to the 5'-end of the target region in the cDNA, with an estimated $T_m$ of 66°C. The JH11 was derived from a portion of the oligo dT-primer adapter; thus, it is specific to the 3'-end of the cDNA with a $T_m$ of 64°C. Both primers were designed to have a recognition site for the restriction enzyme, NotI, at the 5'-end, for use in subsequent cloning of the amplified HCV cDNA.

The PCR reaction was carried out by suspending the cDNA and the primers in 100 microliters of reaction mixture containing the four deoxynucleoside triphosphates, buffer salts and metal ions, and a thermostable DNA polymerase isolated from Thermus aquaticus (Taq polymerase), which are in a Perkin Elmer Cetus PCR kit (N801-0043 or N801-0055). The PCR reaction was perform d for 35 cycles in a Perkin Elmer Cetus DNA thermal cycler. Each cycle consisted of a 1.5 min denaturation step at 94°C, an annealing step at 60°C for 2 min, and a primer extension step at 72°C for 3 min. The PCR products were subjected to Southern blot analysis using a 30 nucleotide probe, JH34, the sequence of which was based upon that of the 3'-terminal region of clone 15e. The sequence of JH34 is:

5′ CTT GAT CTA CCT CCA ATC ATT CAA AGA CTC 3′.

The PCR products detected by the HCV cDNA probe ranged in size from about 50 to about 400 base pairs.

In order to clone the amplified HCV cDNA, the PCR products were cleaved with NotI and size selected by polyacrylamide gel electrophoresis. DNA larger than 300 base pairs was cloned into the NotI site of pUC18S The vector pUC18S is constructed by including a NotI polylinker cloned between the EcoRI and SalI sites of pUC18. The clones were screened for HCV cDNA using th JH34 probe. A number of positive clones were obtained and sequenced. The nucleotide sequence of the HCV cDNA ins rt in one of these clones, 16jh, and the amino acids encoded therein, are shown in Fig. 15. A nucleotide heterogeneity, detected in the sequence of the HCV cDNA in clone 16jh as compared to another clone of this region, is indicated in the figure.

Isolation and Sequence of Clone 6k

Based on the sequence of clone 16jh and clon b5a (see supra), a synthetic probe was made having the following sequence:

5′ TCT TCA ACT GGG CAG TAA GAA CAA AGC TCA 3′.

Screening of the original lambda-gt11 HCV cDNA library (described in E.P.O. Publication No. 318,216) with the probe yielded clones with a frequency of approximately 1 in $10^6$; one of these was called clone 6k (also called C6k), the HCV cDNA sequence of which is shown in Fig. 16. Also shown in the figure are the overlap with clone 16jh, and putative polypeptides encoded within the HCV cDNA. Sequence information on the HCV cDNA in clone 6k was obtained from only one strand. Information on the deposit

of this clon is provided infra, wherein th clone is listed as Lambda gt11 C6k. Confirmation of the C6K s qu nc as part of an ORF ncoding HCV1 polypeptide has be n obtained by sequencing other overlapping clones.

Isolation and Sequence of Clone p131jh

A clone containing sequence from the 3'-region of the HCV genome, and which contains an in-frame stop codon, was isolated essentially as described supra., for the isolation of clones generated by PCR amplification of the 3'region of the genome, except that HCV1 RNA was converted to cDNA using the oligonucleotide

5' AAT TCG CGG CCG CCA TAC GAT TTA GGT GAC ACT ATA GAA $T_{15}$ 3'.

The cDNA was then amplified by the PCR reaction using the primers:

5' TTC GCG GCC GCT ACA GCG GGG GAG ACA T 3'

and

5' AAT TCG CGG CCG CCA TAC GA 3'.

After amplification, the PCR products were precipitated with spermine, digested with NotI, and extracted with phenol. The purified products were cloned into the NotI site of pUC18S, and HCV positive clones were selected using the oligonucleotide:

5' CGA TGA AGG TTG GGG TAA ACA CTC CGG CCT 3'.

The HCV cDNA in one clone, designated p131jh, is shown in Fig. 17. This clone contains an in-frame stop codon for the large ORF contained in the HCV genome.

Isolation and S qu nc of Clone 5'-clone32

A clon containing sequence from th 5'-r gion of the HCV genome, upstream of th sequ nc in clone b114a, was isolated and the nucleotide sequence determined by a modification of the method for the isolation and sequence of clones generated by PCR amplification of the 3'-region of the genome, described in U.S.S.N. 456,637, which is incorporated by reference. Generally, a target region of the genome was amplified by the PCR technique described in Saiki et al. (1986), and in Saiki et al (1988). The HCV RNA which was amplified was obtained by extracting human serum (U.S. clinical isolate, HCV27) using a cold guanidinium thiocyanate method described by Han et al. (1987). The extracted RNA was converted into single stranded cDNA with reverse transcriptase, using a primer, JH94, which is complementary to nucleotides -250 to -223 of the HCV genome (see Fig. 18). The sequence of JH94 is:

5' CCT GCG GCC GCA CGA CAC TCA TAC TAA 3'.

Conversion of single- to double-stranded HCV cDNA was accomplished by tailing the DNA with approximately 20 to 50 dA residues using terminal deoxynucleotidyl transferase (Sambrook et al. (1989), MOLECULAR CLONING), and replicating the tailed molecule using the following oligo-dT primer-adapter, which contains a NotI site, and an sp6 promoter:

| Stuffer | NotI | SP6 Promoter | Primer |
|---|---|---|---|
| AATTC | GCGGCCGC | CATACGATTTAGGTGACACTATAGAA | $T_{15}$ |

The resultant cDNA was subjected to amplification by PCR using two primers, JH94 (described supra.) and JH11, which has the following sequence.

| Primer | S qu nc |
|---|---|
| JH11 (20m r) | AATTCGGGCGGCCGCCATACGA |

The PCR reaction was carried out by suspending the cDNA and the primers in 100 microliters of reaction mixture containing the four deoxynucleoside triphosphates, buffer salts and metal ions, and a thermostable DNA polymerase isolated from Thermus aquaticus (Taq polymerase), which are in a Perkin Elmer Cetus PCR kit (N801-0043 or N801-0055). The PCR reaction was performed for 35 cycles in a Perkin Elmer Cetus DNA thermal cycler. Each cycle consisted of a 1.5 min denaturation step at 94°C, an annealing step at 60°C for 2 min, and a primer extension step at 72°C for 3 min.

The PCR products were digested with NotI, and cloned into pUC18S. Clones containing HCV nucleotide sequences were obtained by screening with a probe, Alex90, which is derived from nucleotides -312 to -283 of the HCV1 genome, and which has the sequence:

5′ ACC ATG AAT CAC TCC CCT GTG AGG AAC TAC 3′.

The HCV cDNAs in the isolated clones were sequenced by the dideoxy chain termination method (Sanger et al. (1977)). The sequence of HCV cDNA in one of the isolated clones, 5′-clone32, spans the region of nucleotides -224 to -341 in Fig. 18.

An analysis of the nucleotide sequence of the HCV cDNA showed that the replicate of the HCV RNA strand contains a GC-rich stretch which may be capable of forming a stable hairpin structure:

```
G-T         A
: :         /\
: :        /  \
A C-C-C-C-C-G C-C-G- 5'
: : : : : : : :   :
: : : : : : : :   :
T-G-G-G-G-G-C-G-A-C- 3'.
```

In the structure, the dashed lines indicate possible hydrogen bonds between complementary nucleotides.

A search in the computer database, Genebank, revealed that homologous sequences were absent from known viral sequences. Thus, this sequence may be unique to th 5'-terminus of the HCV genome.

A hairpin structure may serve as a recognition signal for a transcriptase and/or it may contribute to the stability of the RNA at the 5'-terminus.

<u>Compiled HCV cDNA Sequences</u>

An HCV cDNA sequence has been compiled from a series of overlapping clones derived from various HCV cDNA libraries described herein, and in E.P.O. Publication No. 318,216. The clones from which Fig. 18 has been derived are clone 5'-32, b114a, 18g, ag30a, CA205a, CA290a, CA216a, pi14a, CA167b, CA156e, CA84a, CA59a, K9-1 (also called k9-1), 26j, 13i, 12f, 14i, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g, 39c, 35f, 19g, 26g, 15e, b5a, 16jh, C6k and p131jh. Th methods for isolation of these clones, as well as their sequences, are discussed herein, and in E.P.O. Publication No. 318,216, which is incorporated herein by reference. In Fig. 18, the three dashes above the sequence indicate the position of the putative initiator methionine codon.

Clone b114a overlaps with clones 18g, ag30a, and CA205a, except that clone b114a contains an extra two nucleotides upstream of the sequence in clone 18g (i.e.,

5'-CA). Th se extra two nucleotid s hav b en included in the HCV genomic s quence shown in Fig. 18.

It should be noted that although several of th clones describ d supra. have been obtained from libraries other than the original HCV cDNA lambda-gt11 C library (ATCC No. 40394), these clones contain HCV cDNA sequences which overlap HCV cDNA sequences in the original library. Thus, essentially all of the HCV sequence is derivable from the original lambda-gt11 C library (ATCC No. 40394) which was used to isolate the first HCV cDNA clone (5-1-1). The isolation of clone 5-1-1 is described in E.P.O. Publication No. 318,216, which is incorporated herein by reference.

The putative sequence of the major HCV polyprotein encoded in the composite of HCV1 cDNA is also shown. The first amino acid in the sequence is the putative initiator methionine of the large ORF. The variant amino acids, due to the clonal heterogeneities, are indicated above the sequence. Since the lambda gt11 library was created from serum obtained from one individual (see E.P.O. Publication No. 318,216), the results suggest that variant viral sequences (both nucleotide and amino acid) are present in that individual.

An examination of the composite HCV cDNA sequence shows that besides the large ORF, there are a number of ORFs upstream of that encoding the polyprotein, and within the sequence encoding the polyprotein there are a large number of smaller ORFs in the other two translational frames. The ORFs upstream of the HCV polyprotein are shown in the Table immediately below.

Tabl
ORFs Upstr am of that Encoding th Large
HCV Polyprotein

| Nucl. # | Translation Frame | Amino Acid Sequence |
|---|---|---|
| -310 | 1 | MNHSPVRNYCLHAESV |
| -329 | 3 | MGATLHHESLPCEELL SSRRKRLAMALV |
| -246 | 2 | MSVVQPPGPPLPGEP |
| -127 | 1 | MPGDLGVPPQDC |

The reading frame, position, and size of the ORFs downstream of the sequence encoding the putative initiator MET of the polyprotein are shown in the Table below. The major polyprotein is that translated from reading frame 2.

Table
ORFs Downstream of the Putative Initiator MET Encoding Sequence

| Reading Frame | Size(aa) | Position(bp) |
|---|---|---|
| 1 | 168 | 696 |
| 1 | 105 | 2343 |
| 1 | 119 | 5616 |
| 2 | 3025 | -42 |
| 3 | 160 | 5 |
| 3 | 111 | 1667 |
| 3 | 148 | 6893 |

In addition to the above, an examination of the sequence which is complementary to the genomic strand of HCV RNA also contains several small ORFs. One of these ORFs, which is complementary to nucleotides -341 to +837 in the HCV RNA sequence, encodes a polypeptide of 385 amino acids.

## Comparison of the Sequences of 5'-Regions Obtain d from HCV Isolat s from Different Geographical Locations

Nucl otide sequences from the 5'- regions of HCV isolates from the U.S.A. (HCV18, HCV27), from Italy (HCVI1, HCVI24), and from Korea (HCVK1) were compared.

Isolation of the HCV cDNA sequences was essentially as described supra., for the isolation of 5'-clone32, except for the following. The extracted RNA was reverse-transcribed into cDNA using as primers either JH51 or r16, which are complementary to HCV nucleotides -90 to -73 and 366 to 383, respectively. The sequences of these primers are as follows.

| Primer | Sequence |
|---|---|
| JH51 | 5' CCC AAC ACT ACT CGG CTA 3' |
| r16 | 5' CAC GTA AGG GTA TCG ATG 3' |

Amplification of the HCV dsDNA was by the PCR method using JH93 and JH52 as 5'- and 3'- primers, respectively. Th HCV sequence in JH93 is derived from HCV nucleotides -317 to -296, that in JH52 is from HCV nucleotides -93 to -117; the nucleotide numbers are indicated in parentheses below the sequences. In JH52 the underlined dinucleotide has been mutated to create the NotI site. The sequences of these primers are the following.

| (Primer) | | Stuffer | NotI | HCV sequence | |
|---|---|---|---|---|---|
| (JH93) | 5' | TTC | GCGGCCGC | ACTCCATGAATCACTCCCC | 3' |
| | | | (-317) | | (-296) |
| (JH52) | 5' | AGTCTT | GCGG<u>CC</u>GC | ACGCCCAAATC | 3' |
| | (-93) | | | (-117) | |

After amplificati n, th PCR products were cleaved by NotI, and cloned into pUC18S. The HCV cDNAs were sequenced either by direct s qu ncing aft r amplification by PCR, or alternatively, the cloned HCV cDNAs were sequenced by the primer extension and the dideoxy method. Primer extension and the dideoxy method of sequencing were performed as described supra., for the sequence of 5'-clone32.

The PCR method for direct sequencing used Alex90 (see supra. for the sequence) as the 5'-primer, and r25 as the 3'-primer. Alex90 is derived from HCV nucleotides -312 to -283, and r25 is derived from nucleotides 365 to 342 (See Fig. 18). The sequence of r25 is:

5' ACC TTA CCC AAA TTG CGC GAC CTA 3'.

A comparison of the sequences of the 5'-region of HCV27, HCVK1, HCVI1, HCVI24, and HCV18 with the sequence of the prototype HCV, HCV1, showed the following. The examined 5'- region is highly conserved amongst the 5 HCV isolates. The sequences appeared to be identical except for one nucleotide which was deleted at position -171 in HCVI24, and for the ambiguity in four nucleotides at positions -222 to -219 in isolate HCVK1.

The high levels of sequence conservation in this region may reflect the role of this region in viral replication, and/or transcription, and/or translation.

## Sequence Variations in HCV Isolates from Different Individuals

Isolates of HCV which contain sequences which deviate from CDC/HCV1 were identified in human individuals, some of whom were serologically positive for anti-C100-3 antibodies (EC10 was antibody negative). Identification of these new isolates was accomplished by cloning and sequencing segments of the HCV genome which

had been amplified by the PCR t chnique using CDC/HC1 s qu nc s. Amplification was accomplished essentially based on an HCV/cPCR method. The method utilizes prim rs and probes based upon the HCV cDNA sequences described herein. The first step in the method is the synthesis of a cDNA to either the HCV genome, or its replicative intermediate, using reverse transcriptase. After synthesis of the HCV cDNA, and prior to amplification, the RNA in the sample is degraded by techniques known in the art. A designated segment of the HCV cDNA is then amplified by the use of the appropriate primers. The amplified sequences are cloned, and clones containing the amplified sequences are detected by a probe which is complementary to a sequence lying between the primers, but which does not overlap the primers.

HCV Isolates Isolated from Humans in the U.S.

Blood samples which were used as a source of HCV virions were obtained from the American Red Cross in Charlotte, North Carolina, and from the Community Blood Center of Kansas, Kansas City, Missouri. The samples were screened for antibodies to the HCV C100-3 antigen using an ELISA assay as described in E.P.O. Publication No. 318,216, and subjected to supplemental Western blot analysis using a polyclonal goat anti-human HRP to measur anti-HCV antibodies. Two samples, #23 and #27, from the American Red Cross and from the Community Blood Center of Kansas, respectively, were determined to be HCV positive by these assays.

.Viral particles present in the serum of these samples were isolated by ultracentrifugation under the conditions described by Bradley et al. (1985). RNA was extracted from the particles by digestion with proteinas K and SDS at final concentrations of 10 micrograms/ml proteinase K, and 0.1% SDS; digestion was for 1 hour at 37°C. Viral RNA was further purified by extraction with

chloroform-ph nol, as described in E.P.O. Publication No. 318,216.

HCV RNA in the preparation of RNA was reverse transcribed int cDNA essentially as described in E.P.O. Publication No. 318,216, except that the oligonucleotid JHC 7, which corresponds to the cDNA sequence 1958-1939, and which has the following sequence, was used as primer for the reverse transcriptase reaction.

JHC 7: CCA GCG GTG GCC TGG TAT TG.

After both strands of the cDNA were synthesized, the resulting cDNA was then amplified by the PCR method essentially as described supra. for the isolation of clones generated by PCR amplification, except that the oligonucleotide primers used, i.e., JHC 6 and ALX 80, were designed to amplify a 1080 nucleotide segment of the HCV genome from CDC/HCV1 nucleotides 673 to 1751. The primers, in addition, are designed to incorporate a NOT I restriction site at the 3'-end of the PCR product, and a blunt end at the 5'-terminus. The sequences of the primers is:

ALX 80: TTT GGG TAA GGT CAT CGA TAC CCT TAC GTG;

and

JHC 6: ATA TGC GGC CGC CTT CCG TTG GCA TAA.

ALX 80 corresponds to nucleotides 673-702 of the CDC/HCV1 sequence; JHC 6 corresponds to nucleotides 1752-1738 of the HCV1 (in addition there are 12 extra nucleotides which encode a NotI site). The designation of nucleotides in JHC 6, i.e., a declining number, indicates the placement in the anti-sense strand.

After PCR amplification with the above described primers, the blunt end terminus was converted into a NOT I site as follows. A homopolymer tail of 15 dGs was attached to the PCR product using terminal deoxynucleotide transferase, and the products were again subjected to amplification by PCR using as primers JHC 6 and JHC 13. The latter primer, JHC 13, the sequence of which follows, is designed to contain a NOT I site in addition to an SP6 phage promoter. (The SP6 promoter is described in GENETIC ENGINEERING, J. Setlow Ed. (1988).

JHC 13: AAT TCG CGG CCG CCA TAC GAT TTA GGT GAC
ACT ATA GAA CCC CCC CCC CCC CCC.

In order to clone the amplified HCV cDNA, the PCR products were cleaved with NotI, precipitated with spermine to remove free oligonucleotides (Hoopes et al. (1981)), and cloned into the NotI site of pUC18S (see Section IV.A.34.). The HCV cDNAs in three clones derived from each HCV isolate, were subjected to sequence analysis. Analysis was essentially by the method described in Chen and Seeburg (1985).

Consensus sequences of the clones derived from HCV in samples 23 and 27 are shown in Fig. 46 and Fig. 47, respectively. The variable sequences are also shown in these figures, as are the amino acids encoded in the consensus sequences.

Fig. 39 and Fig. 40 show comparisons of the aligned positive strand nucleotide sequences (Fig. 39) and putative amino acid sequences (Fig. 40) of samples 23, 27, and HCV1. The amino acid sequence of HCV1 in Fig. 39 represents amino acid numbers 129-467 of the HCV polyprotein encoded by the large ORF in the HCV genomic RNA. An examination of Fig. 46 and Fig. 47 show that there are variations in the sequences of the three isolated clones. The sequence variations at the

nucleotide lev 1 and the amino acid level are summarized in the table immediately below. In the table, the polypeptides designated S and NS1 represent amino acid numbers 130 to ~380, and 380 to ~470, respectively. The numbering is from the putative initiator methionine. The terminology S and NS1 is based upon the positioning of the sequences encoding the polypeptides using the Flavivirus model. As discussed above, however, recent evidence suggests that there is not total correlation between HCV and the Flaviviruses with regard to viral polypeptide domains, particularly in the putative E/NS1 domains. Indeed, HCV polypeptides and their coding domains may exhibit substantial deviation from the Flavivirus model.

Table
Sequence Homology

| | Nucleotide Encoding | | | Amino Acid Encoded | | |
|---|---|---|---|---|---|---|
| | overall % | S % | NS1 % | overall % | S % | NS1 % |
| HCV1/HCV23 | 93 | 95 | 91 | 92 | 95 | 87 |
| HCV1/HCV27 | 89 | 93 | 84 | 89 | 95 | 82 |
| HCV23/HCV27 | 89 | 93 | 85 | 90 | 93 | 84 |

Although there are variations in the newly isolated HCV sequences, the cloned sequences from sampl s 23 and 27 (called HCV23 and HCV27) each contain 1019 nucleotides, indicating a lack of deletion and addition mutants in this region in the selected clones. The sequences in Figs. 39 and 40 also show that the isolated sequences are not rearranged in this region.

A comparison of the consensus sequences for HCV1 and for the other isolates of HCV is summarized in the Table, supra. The sequence variations between the chimpanzee isolate HCV1, and the HCVs isolated from humans

are about th sam as that seen between th HCVs of human origin.

It is of interest that the sequence variations in two of the putativ domains is not uniform. The sequence in a putative S region appears to be relatively constant, and randomly scattered throughout the region. In contast, a putative NS1 region has a higher degree of variability than the overall sequence, and the variation appears to be in a hypervariable pocket of about 28 amino acids which is located about 70 amino acids downstream from the putative N-terminus of the putative polyprotein.

Although it may be argued that the detected variations were introduced during the amplification process, it is unlikely that all of the variations are from this result. It has been estimated that Taq polymerase introduces errors into a sequence at approximately one base per 10 kilobases of DNA template per cycle (Saiki et al. (1988)). Based upon this estimate, up to 7 errors may have been introduced during the PCR amplification of the 1019 bp DNA fragment. However, the three subclones of HCV-23 and HCV-27 yielded 29 and 14 base variations, respectively. The following suggest that these variations are naturally occurring. About 60% of the base changes are silent mutations which do not change the amino acid sequence. Variations introduced by the Taq polymerase during PCR amplification would be expected to occur randomly; however, the results show that the variant sequences are clustered in at least one specific region. Moreover, a consensus sequence was derived by sequencing multiple different clones derived from the PCR amplified products.

<u>HCV Isolates from Humans in Italy and in the U.S.</u>

Segments of HCV RNA present in different isolates were amplified by the HCV/cPCR method. These

s gments span a region of ~0.6Kb to ~1.6Kb downstream from the m thionin encoding start codon of th putative HCV polyprotein. The isolat s are from biological specimens obtain d from HCV infect d individuals. More specifically, isolate HCT #18 is from human plasma from an individual in the U.S.A., EC1 and EC10 are from a liver biopsy of an Italian patient, and Th is from a peripheral blood mononucleocyte fraction of an American patient. Comparable segments of HCV RNA have been isolated from a chimpanzee.

RNA was extracted from the human plasma specimens using phenol:CHCl3:isoamyl alcohol extraction. Either 0.1 ml or 0.01 ml of plasma was diluted to a final volume of 1.0 ml, with a TENB/proteinase K/SDS solution (0.05 M Tris-HCL, pH 8.0, 0.001 M EDTA, 0.1 M NaCl, 1 mg/ml Proteinase K, and 0.5% SDS) containing 10 to 40 micrograms/ml polyadenylic acid, and incubated at 37°C for 60 minutes. After this proteinase K digestion, the resultant plasma fractions were deproteinized by extraction with TE (50 mM Tris-HCl, pH 8.0, 1 mM EDTA) saturated phenol, pH 6.5. The phenol phase was separated by centrifugation, and was reextracted with TENB containing 0.1% SDS. The resulting aqueous phases from each extraction were pooled, and extracted twice with an equal volume of phenol/chloroform/isoamyl alcohol [1:1(99:1)], and th n twice with an equal volume of a 99:1 mixture of chloroform/isoamyl alcohol. Following phase separation by centrifugation, the aqueous phase was brought to a final concentration of 0.2 M Na Acetate, and the nucleic acids were precipitated by the addition of two volumes of ethanol. The precipitated nucleic acids were recovered by ultracentrifugation in a SW 41 rotor at 38 K, for 60 minutes at 4°C, or in a microfuge for 10 minutes at 10K, 4°C.

RNA extracted from the liver biopsy was provided by Dr. F. Bonino, Osp dale Maggiore di S. Giovanni Battista, Torino, Italy.

Th mononucleocyte fraction was obtained by sedimentation of the individual's aliquot of blood through Ficoll-Paque® (Pharmacia Corp), using the manufacturer's directions. Total RNA was extracted from the fraction using the guanidinium thiocyanate procedure described in E.P.O. Publication No. 318,216 (See also Choo et al (1989)).

Synthesis of HCV cDNA from the samples was accomplished using reverse transcriptase, and primers derived from clone 156e and from clone K91. These primers, which are anti-sense relative to the genomic RNA, have the following sequences.

156e16B: 5′ CGA CAA GAA AGA CAG A 3′,
and
K91/16B 5′ CGT TGG CAT AAC TGA T 3′.

Following ethanol precipitation, the precipitated RNA or nucleic acid fraction was dried, and resuspended in DEPC treated distilled water. Secondary structures in the nucleic acids were disrupted by heating at 65°C for 10 minutes, and the samples were immediately cooled on ice. cDNA was synthesized using 1 to 3 micrograms of total RNA from liver, or from nucleic acids (or RNA) extracted from 10 to 100 microliters of plasma. The synthesis utilized reverse transcriptase, and was in a 25 microliter reaction, using the protocol specified by the manufacturer, BRL. All reaction mixtures for cDNA synthesis contained 23 units of the RNAase inhibitor, RNASIN™ (Fisher/Promega). Following cDNA synthesis, the reaction mixtures were diluted with water, boiled for 10 minutes, and quickly chilled on ice.

Each s t of sampl s was subjected to two rounds of PCR amplification. Th primers for the reactions wer sel cted to amplify r gions designated "EnvL" and EnvR". The "EnvL" region encompasses nucleotides 669-1243, and putative amino acids 117 to 308; the "EnvR" region encompasses nucleotides 1215-1629, and encodes putative amino acids 300-408 (the putative amino acids are numbered starting from the putative methionine initiation codon). The relationship of these regions relative to the putative polyprotein encoded in the HCV cDNA, and to the polypeptides encoded in the Flavirus model is shown in Fig. 48.

The primers for the first round of PCR reactions were derived from the HCV cDNA sequences in either clone ag30a, clone 156e, or clone k9-1. The primers used for the amplification of the EnvL region were 156e16B (shown supra), and ag30a16A for the sense strand; the amplification of the EnvR region utilized the primer K91/16B (shown supra), and 156e16a for the sense strand. The sequences of the sense strand primers are the following.

For EnvL, ag30a16A: 5′ CTC TAT GGC AAT GAG G 3′,

and

For EnvR, 156e16A: 5′ AGC TTC GAC GTC ACA T 3′ .

The PCR reactions were performed essentially according to the manufacturer's directions (Cetus-Perkin-Elmer), except for the addition of 1 microgram of RNase A. The reactions were carried out in a final volume of 100 microliters. The PCR was perfcrmed for 30 cycles, utilizing a regimen of 94$^{\circ}$C (1 min), 37$^{\circ}$C (2 min), and 72$^{\circ}$C (3 min), with a 7 minute extension at 72$^{\circ}$C for the last cycle. The samples were then extracted with phenol:$CHCl_3$, ethanol precipitated two times, resuspended in 10 mM Tris

HCl, pH 8.0, and concentrated using C ntricon-30 (Amicon) filtration. This procedur efficiently removes olig nucleotides less than 30 nucl otides in size; thus, th primers fr m the first round of PCR amplification ar removed.

The Centricon-30 concentrated samples were then subjected to a second round of PCR amplification using probes designed from clones 202a and 156e for the EnvL region, and from 156e and 59a for the EnvR region. The primers for amplification of the EnvL region have the following sequences.

202aEnv41a: 5′ CTT GAA TTC GCA ATT TGG GTA
AGG TCA TCG ATA CCC TTA CG 3′

and

156e38B′: 5′ CTT GAA TTC GAT AGA GCA ATT
GCA ACC TTG CGT CGT CC 3′.

The primers for amplification of the EnvR region in RNAs derived from humans have the following sequences.

156e38A′: 5′ CTT GAA TTC GGA CGA CGC AAG
GTT GCA ATT GCT CTA TC 3′

and

59aEnv39C: 5′ CTT GAA TTC CAG CCG GTG TTG
AGG CTA TCA TTG CAG TTC 3′.

Amplification by PCR was for 35 cycles utilizing a regimen of 94°C (1 min), 60°C (1 min), and 72°C (2 min), with a 7 minute extension at 72°C for the last cycle. The samples were then extracted with phenol:$CHCl_3$, precipitated two times, and digested with EcoRI. The PCR reaction products

were analyzed by separation of the products by lectrophor sis on 6% polyacrylamide gels. DNA of approximat ly the stimated size of the expect d PCR product was electroeluted from the gels, and subcloned into either a pGEM-4 plasmid vector or into lambda gt11. The expected product sizes for the EnvL and EnvR after the first round of amplification are 615 bp and 683 bp, respectively; after the second round of amplification the expected product sizes for EnvL and EnvR are 414 bp and 575 bp, respectively. The plasmids containing the amplified products were used to transform host cells; the pGEM-4 plasmid was used to transform DH5-alpha, and lambda gt11 was used to transform C600 delta-HFL. Clones of the transformed cells which either hybridized to the appropriate HCV probes (described below), or those which had inserts of the correct size were selected. The inserts were then cloned in M13 and sequenced.

The probes for all of the HCV/cPCR products consisted of $^{32}P$ labeled sections of HCV cDNA which had been prepared by PCR amplification of a region of clone 216 (using CA216a16A and 216a16B as primers), and of clone 84 (using CA84a16A and CA84a16B or CA84a16C as primers); $^{32}P$ was introduced into the PCR products by nick translation. The probes for the first and second round of EnvL amplification were from clone 216. Those for the first round of EnvR amplification were from 84 (i.e., CA84a16A and CA84a16B), for the second round of EnvL amplification were CA84a16A and CA84a16C. These probes did not overlap the primers used in the HCV/cPCR reactions. The sequence of the primers for the PCR amplification of the probes is in the following table.

Table

| Primer | Clon | Sequence |
|---|---|---|
| CA216a16A | 216 | 5′ TGA ACT ATG CAA CAG G 3′ |
| CA216a16B | 216 | 5′ GGA GTG TGC AGG ATG G 3′ |
| CA84a16A | 84 | 5′ AAG GTT GCA ATT GCT C 3′ |
| CA84a16B | 84 | 5′ ACT AAC AGG ACC TTC G 3′ |
| CA84a16C | 84 | 5′ TAA CGG GTC ACC GCA T 3′ |

Sequence information on variants in the EnvL region was obtained from 3 clones from HCT #18, 2 clones from TH, 3 clones from EC1, and from the HCV1 clones described in E.P.O. Publication No. 318,216, and supra. A comparison of the composite nucleotide sequence of each isolate derived from these clones is shown in Fig. 49. In the figure, each sequence is shown 5′ to 3′ for the sense strand for the EnvL region, and the sequences have been aligned. The vertical lines and capital letters indicate sequence homology, the absence of a line and an uncapitalized letter indicates a lack of homology. The sequences shown in the lines are as follows: line 1, Thorn; line 2, EC1; line 3, HCT #18; line 4, HCV1.

Sequence information on variants in the EnvR region was obtained from two clones of EC10, and from th HCV1 clones described in E.P.O. Publication No. 318,216 and supra.. The two EC10 clones differed by only one nucleotide. A comparison of the nucleotide sequences of EC10(clone 2) and a composite of the HCV1 sequences is shown in Fig. 50; each sequence is shown 5′ to 3′ for the sense strand of the EnvR region, and the sequences have been aligned. The double dots between the sequences indicate sequence homology.

A comparison of the amino acid sequences encoded in the EnvL (amino acids #117-308) and EnvR region (amino acids #300-438) for each of the isolates is shown in Fig.

51 and Fig. 52, respectively. Included in the Figures are s quences for the isolat s JH23 and JH27, describ d supra. Also indicated are sequenc s from a Japanese isolate; these sequences were provided by Dr. T. Miyamura, Japan. In the figures, the amino acid sequence for the region is given in its entirety for HCV1, and the non-homologous amino acids in the various isolates are indicated.

As seen in Fig. 51, In the EnvL region there is overall about a 93% homology between HCV1 and the other isolates. HCT18, Th, and EC1 have about a 97% homology with HCV1; JH23 and JH27 have about 96% and about 95% homology, respectively, with HCV1. Fig. 52 shows that th homologies in the EnvR region are significantly less than in the EnvL region; moreover, one subregion appears to be hypervariable (i.e., from amino acid 383-405). This data is summarized in the Table immediately below.

Table
Homology of EnvR Region

| Isolate | Percent Homology with HCV1 | |
|---|---|---|
| | AA330-AA438 | AA383-AA405 |
| JH23(U.S.) | 83 | 57 |
| JH27(U.S.) | 80 | 39 |
| Japanese | 73 | 48 |
| EC10 (Italy) | 84 | 48 |

Detection of Positive and Negative Strand 5'-HCV RNA in Serum

The RNA in HCV27, isolated from serum, was analyzed for the presence of positive and negative strands using the PCR method. The PCR method was performed essentially as described above, except for the following. The extracted HCV27 RNA was reverse transcribed into single-stranded cDNA using as a primer either Alex90 or JH52 (see supra. for the sequences). The sequence of

Alex90 matches that in nucleotides -312 to -283 of the positiv strand of HCV RNA, wh reas JH52 matches that of nucl otid s -117 to -93 of the negativ strand. the r sulting single-stranded HCV cDNAs were ach separately amplified by PCR using Alex90 and JH52. Detection of the amplified products was accomplished by Southern blotting, using Alex89 as the probe. Alex89 matches nucleotide numbers -203 to -175 of HCV RNA. The sequence of Alex89 is:

5′ CCA TAG TGG TCT GCG GAA CCG GTG AGT ACA 3′.

The analysis indicated that, by this method, the signals of the amplified products of both RNA strands were of equal intensity. These results are suggestive that HCV RNA in the 5′-region may exist as double-stranded RNA.

Probes for Sandwich Hybridization for HCV

This example exemplifies the sets of label and capture probes useful to detect HCV RNA in biological samples, using essentially the assay described in U.S. Patent No. 4,868,105. The method is a solution-phase sandwich hybridization assay which utilizes both capture and label probes which hybridize to target sequences in an analyte nucleic acid. In the screening of biological samples for HCV, the probes used bind to conserved regions of the HCV genome, and the HCV binding regions are selected for their uniqueness to the HCV genome. The regions which bind to the binding partner of the capture probe, or the portion of the label probe which binds to the labeling moiety (or to an amplifying multimer if the method described in E.P.O. Publication No. 317,077 is used), are selected such that they do not bind to any of the known sequences in the databank or in HCV, and which have the appropriate content of Gs and Cs to allow stable duplex formation with their complements under the selec-

tion conditions. Th capture and label probes are in sets, and th probes of one set do not intersperse with the probes of another set. These probes are comprised of sequences which are complementary to the following nucleotide sequences in the coding strand of the prototype HCV cDNA sequence shown in Fig. 18.

Set 1

| Probe type | Probe Number | Complement of Nucleotide Numbers |
|---|---|---|
| Capture | 42.XT1.1 | -318 to -289 |
| Capture | 42.XT1.2 | -285 to -256 |
| Capture | 42.XT1.3 | -252 to -223 |
| Capture | 42.XT1.4 | -219 to -190 |
| Label | 42.LLA2C.5 | -186 to -157 |
| Label | 42.LLA2C.6 | -153 to -124 |
| Label | 42.LLA2C.7 | -120 to -91 |
| Label | 42.LLA2C.8 | -87 to -58 |
| Label | 42.LLA2C.9 | -54 to -25 |
| Label | 42.LLA2C.10 | -21 to 9 |
| Label | 42.LLA2C.11 | 13 to 42 |
| Label | 42.LLA2C.12 | 46 to 75 |
| Label | 42.LLA2C.13 | 79 to 108 |
| Label | 42.LLA2C.14 | 112 to 141 |
| Label | 42.LLA2C.15 | 145 to 174 |

Set 2

| Probe typ | Probe Number | Complement of Nucleotide Numbers |
|---|---|---|
| Capture | 42.16.XT1 | 4378 to 4407 |
| Capture | 42.17.XT1 | 4411 to 4440 |
| Capture | 42.18.XT1 | 4444 to 4473 |
| Capture | 42.19.XT1 | 4477 to 4506 |
| Capture | 42.20.XT1 | 4510 to 4539 |
| Label | 42.21.LLA2C | 4543 to 4572 |
| Label | 42.22.LLA2C | 4576 to 4605 |
| Label | 42.23.LLA2C | 4609 to 4638 |
| Label | 42.24.LLA2C | 4642 to 4671 |
| Label | 42.25.LLA2C | 4675 to 4704 |
| Label | 42.26.LLA2C | 4708 to 4737 |
| Label | 42.27.LLA2C | 4771 to 4770 |
| Label | 42.28.LLA2C | 4774 to 4803 |
| Label | 42.29.LLA2C | 4807 to 4836 |
| Label | 42.30.LLA2C | 4840 to 4869 |
| Label | 42.31.LLA2C | 4873 to 4902 |

S t 3

| Probe type | Probe Number | Complement of Nucleotide Numbers |
|---|---|---|
| Capture | 42.32.XT1 | 4056 to 4085 |
| Capture | 42.33.XT1 | 4089 to 4085 |
| Capture | 42.34.XT1 | 4122 to 4151 |
| Capture | 42.35.XT1 | 4155 to 4184 |
| Label | 42.36.LLA2C | 4188 to 4217 |
| Label | 42.37.LLA2C | 4221 to 4250 |
| Label | 42.38.LLA2C | 4254 to 4283 |
| Label | 42.39.LLA2C | 4287 to 4316 |
| Label | 42.40.LLA2C | 4230 to 4349 |
| Label | 42.41.LLA2C | 4353 to 4382 |
| Label | 42.42.LLA2C | 4386 to 4415 |
| Label | 42.43.LLA2C | 4419 to 4448 |

In the above sets, each capture probe contains, in addition to the sequences complementary to the HCV sequences, the following sequence downstream of the HCV sequence (i.e., at the 3'-end):

5' CTT CTT TGG AGA AAG TGG TG 3'.

The sequence common to each capture probe is complementary to a sequence in the binding partner(s), so that after hybridization, the duplex can be captured via affixation to the solid phase.

Also, in each set, each label probe contains, in addition to the sequences complementary to the HCV sequences, the following sequence downstream of the HCV sequence:

5' TTA GGC ATA GGA CCC GTG TC 3'.

If the meth d described in E.P.O. Publication No. 317,077 is us d, th sequenc common to ach lab l probe is complementary to a sequence in a multimer, to allow hybrid duplex formation with that multimer.

The sequences of the probes in the above sets are shown in Fig. 19.

## Detection of HCV Polynucleotide Sequences Using PCR Amplification

In the generalized method for amplification of HCV RNA by cPCR it is contemplated that the RNA strand is a virion or mRNA strand, which is a "sense" strand. However, it is also possible that replicative intermediate forms may also be detected which would be "anti-sense"; in this case the primer would be "sense". An RNA sense strand containing the target region is hybridized with an anti-sense primer which primes the synthesis of the replicate strand containing the target. cDNA to the RNA template is synthesized with a primer- and template-dependent reverse transcriptase. The cDNA in the resulting RNA:cDNA hybrid is released by denaturation and treatment with RNAse. Primers are annealed to the cDNA, and extended with a primer- and template-dependent DNA polymerase. The products are denatured, re-annealed to primers, and a second round of synthesis is conducted. A number of cycles are run until the amplified product containing the target region is in a desired amount, which is at least a detectable level.

## Detection of Amplified HCV Nucleic Acid Sequences derived from HCV Nucleic Acid Sequences in Liver and Plasma Specimens from Chimpanzees with NANBH

HCV nucleic acids present in liver and plasma of chimpanzees with NANBH, and not in control chimpanzees, were amplified using essentially the polymerase chain re-

acti n (PCR) t chnique described by Saiki et al. (1986). The primer oligonucleotides were derived from the HCV cDNA s qu nces in clone 81 (Fig. 22), or clones 36 (Fig. 23) and 37b (Fig. 24). The amplified sequences were detected by gel electrophoresis and a modified Southern blotting method, using as probes the appropriate cDNA oligomer or nick-translated cDNA sequence with a sequence from the region between, but not including, the two primers.

Samples of RNA containing HCV sequences to be examined by the amplification system were isolated from liver biopsies of three chimpanzees with NANBH, and from two control chimpanzees. The isolation of the poly $A^+$ RNA fraction was by the guanidinium thiocyanate procedure described in Maniatis et al. (1982).

Samples of RNA which were to be examined by the amplification system were also isolated from the plasmas of two chimpanzees with NANBH, and from one control chimpanzee, as well as from a pool of plasmas from control chimpanzees. One infected chimpanzee had a titer equal to or greater than $10^6$ CID/ml, and the other infected chimpanzee had a titer equal to or greater than $10^5$ CID/ml.

The nucleic acids were extracted from the plasma as follows. Either 0.1 ml or 0.01 ml of plasma was diluted to a final volume of 1.0 ml, with a TENB/proteinase K/SDS solution (0.05 M Tris-HCL, pH 8.0, 0.001 M EDTA, 0.1 M NaCl, 1 mg/ml Proteinase K, and 0.5% SDS) containing 10 micrograms/ml polyadenylic acid, and incubated at 37°C for 60 minutes. After this proteinase K digestion, the resultant plasma fractions were deproteinized by extraction with TE (10.0 mM Tris-HCl, pH 8.0, 1 mM EDTA) saturated phenol. The phenol phase was separated by centrifugation, and was reextracted with TENB containing 0.1% SDS. The resulting aqueous phases from each extraction were pooled, and extracted twice with an equal volume of phenol/chloroform/isoamyl alcohol

[1:1(99:1)], and th n twice with an equal volume of a 99:1 mixture of chloroform/isoamyl alcohol. Following phase s paration by centrifugation, the aqueous phase was brought to a final concentration of 0.2 M Na Acetate, and the nucleic acids were precipitated by the addition of two volumes of ethanol. The precipitated nucleic acids were recovered by ultracentrifugation in a SW 41 rotor at 38 K, for 60 minutes at 4°C.

In addition to the above, the high titer chimpanzee plasma and the pooled control plasma alternatively were extracted with 50 micrograms of poly A carrier by the procedure of Chomcyzski and Sacchi (1987). This procedure uses an acid guanidinium thiocyanate extraction. RNA was recovered by centrifugation at 10,000 RPM for 10 minutes at 4°C in an Eppendorf microfuge.

On two occasions, prior to the synthesis of cDNA in the PCR reaction, the nucleic acids extracted from plasma by the proteinase K/SDS/phenol method were further purified by binding to and elution from S and S Elutip-R Columns. The procedure followed was according to the manufacturer's directions.

The cDNA used as a template for the PCR reaction was derived from the nucleic acids (either total nucleic acids or RNA) prepared as described above. Following ethanol precipitation, the precipitated nucleic acids were dried, and resuspended in DEPC treated distilled water. Secondary structures in the nucleic acids were disrupted by heating at 65°C for 10 minutes, and the samples were immediately cooled on ice. cDNA was synthesized using 1 to 3 micrograms of total chimpanzee RNA from liver, or from nucleic acids (or RNA) extracted from 10 to 100 microliters of plasma. The synthesis utilized reverse transcriptase, and was in a 25 microliter reaction, using the protocol specified by the manufacturer, BRL. The primers for cDNA synthesis were those also utilized in the PCR reaction, described below. All reaction mixtures for

cDNA synth sis contain d 23 units of th RNAase inhibitor, RNASIN® (Fish r/Prom ga). Following cDNA synthesis, th reaction mixtures were diluted with wat r, boiled for 10 minutes, and quickly chilled on ice.

The PCR reactions were performed essentially according to the manufacturer's directions (Cetus-Perkin-Elmer), except for the addition of 1 microgram of RNase A. The reactions were carried out in a final volume of 100 microliters. The PCR was performed for 35 cycles, utilizing a regimen of 37°C (2 min), 72°C (3 min), and 94°C (1 min).

The primers for cDNA synthesis and for the PCR reactions were derived from the HCV cDNA sequences in either clone 81, clone 36, or clone 37b. (The HCV cDNA sequences of clones 81, 36, and 37b are shown in Figs. 22, 23, and 24, respectively.) The sequences of the two 16-mer primers derived from clone 81 were:

5' CAA TCA TAC CTG ACA G 3'
and
5' GAT AAC CTC TGC CTG A 3'.

The sequence of the primer from clone 36 was:

5' GCA TGT CAT GAT GTA T 3'.

The sequence of the primer from clone 37b was:

5' ACA ATA CGT GTG TCA C 3'.

In the PCR reactions, the primer pairs consisted of either the two 16-mers derived from clone 81, or the 16-mer from clone 36 and the 16-mer from clone 37b.

The PCR reaction products were analyzed by separation of the products by alkaline gel electrophoresis, followed by Southern blotting, and detec-

ti n f th amplifi d HCV-cDNA sequences with a $^{32}P$-label d int rnal oligonucleotide probe derived from a r gion of the HCV cDNA which do s not overlap th prim rs. Th PCR reaction mixtures were extracted with phenol/ chloroform, and the nucleic acids precipitated from the aqueous phase with salt and ethanol. The precipitated nucleic acids were collected by centrifugation, and dissolved in distilled water. Aliquots of the samples were subjected to electrophoresis on 1.8% alkaline agarose gels. Single stranded DNA of 60, 108, and 161 nucleotide lengths were co-electrophoresed on the gels as molecular weight markers. After electrophoresis, the DNAs in the gel were transferred onto Biorad Zeta Probe™ paper. Prehybridization and hybridization, and wash conditions were those specified by the manufacturer (Biorad).

The probes used for the hybridization-detection of amplified HCV cDNA sequences were the following. When the pair of PCR primers were derived from clone 81, the probe was an 108-mer with a sequence corresponding to that which is located in the region between the sequences of the two primers. When the pair of PCR primers were derived from clones 36 and 37b, the probe was the nick-translated HCV cDNA insert derived from clone 35, the nucleotide sequence of which is shown in Fig. 34. The primers are derived from nucleotides 155-170 of the clone 37b insert, and 206-268 of the clone 36 insert. The 3'-end of the HCV cDNA insert in clone 35 overlaps nucleotides 1-186 of the insert in clone 36; and the 5'-end of clone 35 insert overlaps nucleotides 207-269 of the insert in clone 37b. (Compare Figs. 23, 34 and 24.) Thus, the cDNA insert in clone 35 spans part of the region between the sequences of the clone 36 and 37b derived primers, and is useful as a probe for the amplified sequences which include these primers.

Analysis of the RNA from the liver specimens was according to the above procedure utilizing both sets of

primers and probes. The RNA from the liver of the three chimpanzees with NANBH yielded positive hybridization results for amplificati n sequences of the expected size (161 and 586 nucleotides for 81 and 36 and 37b, respectively), while the control chimpanzees yielded negative hybridization results. The same results were achieved when the experiment was repeated three times.

Analysis of the nucleic acids and RNA from plasma was also according to the above procedure utilizing the primers and probe from clone 81. The plasmas were from two chimpanzees with NANBH, from a control chimpanzee, and pooled plasmas from control chimpanzees. Both of the NANBH plasmas contained nucleic acids/RNA which yielded positive results in the PCR amplified assay, while both of the control plasmas yielded negative results. These results have been repeatedly obtained several times.

Defective viruses have been known to occur in RNA viruses. By using PCR technology it is possible to design primers to amplify sequences of the HCV genome. By analysis of the amplified products, it is expected to be able to identify both defective versions of the viral genome as well as wild-type viral species. Accordingly, using two primers based on known HCV sequence, one can predict accurately the expected size of the PCR product. Any larger species observed by gel electrophoresis and hybridization analysis could represent potential variant genomes. Alternatively, any smaller species observed in this fashion might represent defective agents. Analyses of these types would be useful in confirming the exact origin of the known HCV sequence, whether it is indeed a wild-type viral sequence or a defective genome. Techniques and methods for these analyses are well known in the art and have been previously described. This methodology will enable one skilled in the art to obtain

related (wild-type or defective) forms of the viral genom .

D t ction of Sequ nces in Captured Particles
Which When Amplified by PCR
Hybridize to HCV cDNA Derived from Clone 81

The RNA in captured particles was obtained as described below. The analysis for sequences which hybridize to the HCV cDNA derived from clone 81 was carried out utilizing the PCR amplification procedure, as described supra., except that the hybridization probe was a kinased oligonucleotide derived from the clone 81 cDNA sequence. The results showed that the amplified sequences hybridized with the HCV cDNA probe.

Particles were captured from HCV infected chimpanzee plasma using polystyrene beads coated with an immunopurified antibody directed against the polypeptide encoded in clone 5-1-1. The procedure for producing th immunopurified antibody preparation is described in E.P.O. Publication No. 318,216, which is commonly owned by the herein assignee, and which is incorporated herein by reference. Briefly, the HCV polypeptide encoded within clone 5-1-1 was expressed as a fusion polypeptide with superoxide dismutase (SOD). This was accomplished by subcloning the clone 5-1-1 cDNA insert into the expression vector pSODcf1 (Steimer et al. (1986)). DNA isolated from pSODcf1 was treated with BamHI and EcoRI, and the following linker was ligated into the linear DNA created by the restriction enzymes:

```
5'  GAT CCT GGA ATT CTG ATA AGA
        CCT TAA GAC TAT TTT AA  3'
```

After cloning, the plasmid containing the insert was isolated. Plasmid containing the insert was restricted with EcoRI. The HCV cDNA insert in clone 5-1-1 was

excised with EcoRI, and ligated into this EcoRI lin arized plasmid DNA. Th DNA mixture was used to transform E. coli strain D1210 (Sadl r et al. (1980)). Recombinants with the 5-1-1 cDNA in the correct orientation for expression of the ORF were identified by restriction mapping and nucleotide sequencing. Recombinant bacteria from one clone were induced to express the SOD-$NANB_{5-1-1}$ polypeptide by growing the bacteria in the presence of IPTG. The fusion polypeptide was purified from the recombinant E. coli by differential extraction of the cell extracts with urea, followed by chromatography on anion and cation exchange columns. The purified SOD-$NANB_{5-1-1}$ polypeptide was attached to a nitrocellulose membrane. Antibody in samples of HCV infected serum was absorbed to the matrix-bound polypeptide. After washing to remove non-specifically bound materials and unbound materials, the bound antibody was released from the bound polypeptide.

## cPCR Method to Detect HCV RNA in Liver and in Serum from Individuals with NANBH.

The reliability and utility of a modified form of the PCR assay, i.e., a cPCR assay, for detecting HCV infection was determined by performing the assay on total liver RNA and on serum from infected individuals. In th cPCR assay, putative viral RNA in the sample is reverse transcribed into cDNA with reverse transcriptase; a segment of the resulting cDNA is then amplified utilizing a modified version of the PCR technique described by Saiki et al. (1986). The primers for the cPCR technique are derived from HCV RNA, which can be identified by the family of HCV cDNAs provided herein. Amplified product corresponding to the HCV-RNA is detected utilizing a probe derived from the family of HCV cDNAs provided herein.

The cPCR/HCV assay used in these studies were performed utilizing the following methods for the prepara-

tion f RNA, the reverse transcription of the RNA into cDNA, th amplification of sp cific s gments of the cDNA by PCR, and the analysis of the PCR products.

RNA was extracted from liver utilizing th guanidium isothiocyanate method for preparing total RNA described in Maniatis et al. (1982).

In order to isolate total RNA from plasma, the plasma was diluted five- to ten-fold with TENB (0.1 M NaCl, 50 mM Tris-HCl, pH 8.0, 1 mM EDTA) and incubated in a Proteinase K/SDS solution (0.5% SDS, 1 mg/ml Proteinase K, 20 micrograms/ml Poly A carrier) for 60 to 90 minutes at $37^{o}C$. The samples were extracted once with phenol (pH 6.5), the resulting organic phase was re-extracted once with TENB containing 0.1% SDS, and the aqueous phases of both extractions were pooled and extracted twice with an equal volume of phenol/$CHCl_3$/isoamyl alcohol [1:1(99:1)]. The resulting aqueous phases were extracted with an equal volume of $ChCl_3$/isoamyl alcohol (99:1) twice, and ethanol precipitated using 0.2 M sodium acetate, pH 6.5, and 2.5 volumes of 100% ethanol; precipitation was overnight at -$20^{o}C$.

The cDNA used as a template for the PCR reaction was prepared utilizing the designated samples for preparation of the corresponding cDNAs. Each RNA sample (containing either 2 micrograms of heat denatured total chimpanzee liver RNA, RNA from 2 microliters of plasma, or 10% of the RNA extracted from 10mm X 4 mm cylindrical human liver biopsies) was incubated in a 25 microliter reaction containing 1 micromolar of each primer, 1 millimolar of each deoxyribonucleotide triphosphate (dNTP), 50 millimolar Tris-HCL, pH 8.3, 5 millimolar $MgCl_2$, 5 millimolar dithiothreitol (DTT), 73 millimolar KCl, 40 units of RNase inhibitor (RNASIN), and 5 units of AMV reverse transcriptase. The incubation was for 60 minutes at $37^{o}C$. Following cDNA synthesis, the reactions

were diluted with 50 microliters of deionized water (DIW), boiled for 10 minutes, and cooled on ice.

Amplification of a segment of the HCV cDNA was p rformed utilizing two synthetic oligomer 16-mer primers whose sequences were derived from HCV cDNA clones 36 (anti-sense) and 37b (sense). The sequence of the primer from clone 36 was:

5′ GCA TGT CAT GAT GTA T 3′.

The sequence of the primer from clone 37b was:

5′ ACA ATA CGT GTG TCA C 3′.

The primers were used at a final concentration of 1 micromolar each. In order to amplify the segment of HCV cDNA which is flanked by the primers, the cDNA samples were incubated with 0.1 microgram of RNAse A and the PCR reactants of the Perkin Elmer Cetus PCR kit (N801-0043 or N801-0055) according to the manufacturer's instructions. The PCR reaction was performed for either 30 cycles or 60 cycles in a Perkin Elmer Cetus DNA thermal cycler. Each cycle consisted of a 1 minute denaturation step at 94°C, an annealing step of 2 minutes at 37°C, and an extension step of 3 minutes at 72°C. However, the extension step in the final cycle (30 or 60) was 7 minutes rather than 3 minutes. After amplification the samples were extracted with an equal volume of phenol: chloroform (1:1), followed by extraction with an equal volume of chloroform, and then the samples were precipitated with ethanol containing 0.2 M sodium acetate.

The cPCR products were analyzed as follows. The products were subjected to electrophoresis on 1.8% alkaline agarose gels according to Murakawa et al. (1988), and transferred onto Zeta™ Probe paper (BioRad Corp.) by blotting gels overnight in 0.4 M NaOH. The blots were

neutraliz d in 2 X SSC (1 X SSC contains 0.15 M NaCl, 0.015 M sodium citrat ), pr hybridized in 0.3 M NaCl, 15 mM sodium phosphag buffer, pH 6.8, 15 mM EDTA, 1.0% SDS, 0.5% nonfat milk (Carnation Co.), and 0.5 mg/ml sonicated denatured salmon sperm DNA. The blots to be analyzed for HCV cDNA fragments were hybridized to a $^{32}P$-labeled probe generated by nick translation of the HCV cDNA insert sequence in clone 35, described in E.P.O. Publication No. 318,216. After hybridization, the blots were washed in 0.1 X SSC (1 X SSC contains 0.15M NaCl, 0.01M Na citrate) at 65°C, dried, and autoradiographed. The expected product size is 586 nucleotides in length; products which hybridized with the probe and migrated in the gels in this size range were scored as positive for viral RNA.

As a control, cPCR primers designed to amplify alpha-1 anti-trypsin mRNA was performed to verify the presence of RNA in each sample analyzed. The coding region of the alpha-1 anti-trypsin gene is described in Rosenberg et al. (1984). Synthetic oligomer 16-mer primers designed to amplify a 365 nucleotide fragment of the coding region of the alpha-1 antitrypsin gene were deriv d from nucleotides 22-37 (sense) and nucleotides 372-387 (antisense). The PCR products were detected using a $^{32}P$ nick-translated probe which lies between, and not including, the cDNA/PCR primer sequences.

Due to the extreme sensitivity of the PCR reaction, all samples were run a minimum of three times. All false positive signals were eliminated when the following precautions were taken: 1) eliminating aerosols by using screw capped tubes with rubber O-ring seals; 2) pipetting with Ranin Microman™ positive displacement pipetters with disposable pistons/capillaries; and 3) selecting the oligonucleotide sequences for the cDNA and PCR primers from two non-contiguous cDNA clones.

D tection of HCV RNA in Liver Samples by a cPCR M thod

The cPCR assay was perform d on total RNA isolated from livers of three chimpanzees experimentally infected with a NANBH agent, and from liver biopsies of Italian patients diagnosed as having chronic NANBH.

Fig. 25A shows the results of the cPCR assay using 1 microgram of each preparation of total liver RNA. The RNA was isolated from liver samples of a chimpanzee in the chronic phase of NANBH (910)(lane 1), two chimpanzees in the acute phase of infection (1028 and 508)(lanes 2 and 3, respectively). PCR was performed on the samples in lanes 1-3 for 30 cycles and the autoradiogram of the blot containing those lanes was exposed for 5 hours. cDNA from 1 microgram of total RNA from acutely infected animal 1028 (lane 4), and three uninfected chimpanzees (lanes 5-7), were amplified for 60 cycles and the autoradiograms containing those lanes were exposed for 7 days. $^{32}P$ labeled MspI-digested pBR322 DNA served as markers on all the autoradiograms. It may be seen from the results that cDNA corresponding to HCV RNA was seen only in the samples from chimpanzees with NANBH, whether acute or chronic (lanes 1, 3, and 4). The cPCR products in these lanes migrated between marker fragments of 527 and 622 nucleotides (not shown).

Fig. 25B shows the results of the cPCR assay using 10% of the RNA extracted from 10mm X 4mm liver biopsy cylinders from 15 chronic NANB patients (lanes 1-15), one patient with cryptogenic liver disease (lane 16) and one control sample from a patient with chronic Hepatitis B (lane 17). Amplification by PCR was for 30 cycles and the autoradiogram for the blots were exposed for 4 days, except that lane 1 was exposed for 15 hours. As seen from the results, 9/15 (60%) of the human samples were positive for HCV RNA (lanes 1,2,4,6,7,10-13). One patient diagnosed with cryptogenic liver disease (lane 16)

and on pati nt with a chronic HBV inf ction (lane 17) were repeat dly negative in the cPCR assay.

Comparison of the HCV/cPCR Assay on Human Liver Biopsies and RIA of Serum Using HCV C100-3 Polypeptide

SOD/HCV C100-3 polypeptide (also called C100) is a recombinant fusion polypeptide which contains 363 viral amino acids. The polypeptide is useful for detecting antibodies to HCV (See Kuo et al. (1989)). The method for preparing C100 is described in E.P.O. Publication No. 318,216.

Radioimmune assay using C100 was performed on the sera collected from the same 17 human patients whose liver samples were subjected to HCV/cPCR assay as described supra. The sera was collected on the same day as the liver biopsies. The assay was performed essentially as described in E.P.O. Publication No. 318,216, which is commonly owned and incorporated herein by reference. Briefly, Microtiter plates (Immulon 2, Removeawell strips) were coated with 0.1 microgram of purified C100. The coated plates were incubated for 1 hour at 37°C with the serum samples (100 microliters of a 1:100 dilution) or appropriate controls. After incubation, the unbound material was removed, the plates were washed, and complexes of human antibody-C100 were detected by incubation with $^{125}$I-labeled sheep anti-human immunoglobulin. Unbound labeled antibody was removed by aspiration, and the plates were washed. The radioactivity in individual wells was determined.

.The results of the RIA showed that sixty-seven percent of these samples were positive for anti-C100 antibodies. Sera from the patient diagnosed with cryptogenic liver disease was positive for anti-C100 antibodies, although the levels of viral RNA were undetectable in the patient's liver in this sample. The level of correlation between the presence of anti-C100 antibodies and HCV RNA

was s venty p rc nt; two patients who wer negative for antibodies by RIA had significant levels of HCV RNA in their livers (data not shown).

The results indicate that virus is frequently present in the liver of patients with circulating anti-C100 antibodies, and confirms claims that the presence of anti-C100 antibodies accurately reflects exposure to HCV. Moreover, taken together, these results indicate that HCV of this type accounts for NANBH in at least 75% of the patients in this study, and that the predominant strain of HCV in Italy appears to be closely related to the strain of HCV prevalent in the United States.

HCV/cPCR Assay of Sera: Detection of Viral RNA in Acute Phase Infection in Chimpanzees

The temporal relationship between the display of liver damage, the presence of HCV RNA, and the presence of anti-HCV antibodies was monitored in serum from two experimentally infected chimpanzees with NANBH (nos. 771 and 910). Liver damage was determined by alanine amino transferase (ALT) levels; the presence of HCV RNA was determined by the HCV cPCR assay described above; anti-HCV antibodies were detected utilizing the C100 RIA.

The HCV/cPCR analysis was performed on RNA extracted from 1 microliter of chimpanzee plasma. Serum was taken from chimpanzee 771 on days 25, 32, 70 and 88 post-infection; cPCR was performed for 30 cycles and the autoradiogram was exposed for 18 days. Serum was taken from chimpanzee 910 on days 11, 28, and 67 post-infection; cPCR was performed for 60 cycles and the autoradiogram was exposed for 5 days.

The results of the assays are shown in Fig. 26A for chimpanzee 771, and Fig. 26B for chimpanzee 910. From a comparison of Figs. 26A and 26B, it appears that an early, well defined peak of ALT values during acute

hepatitis correlates with th presenc of viral RNA in the inf ct d individual.

The data also indicate that th presence of HCV RNA, which is indicativ of a state of viremia, precedes the presence of anti-HCV antibodies. Chimpanzee 771 (Fig. 26A) exhibited a clearly defined acute episode of post-transfusion NANBH at 28 days, as characterized by an initial peak of ALT levels. HCV RNA was detected in th serum collected at day 25, and at day 32. However, during this acute phase, anti-HCV antibodies were absent. In contrast, at day 70 HCV RNA was below the experimental level of detection, and anti-HCV antibodies were rising. At day 88, HCV RNA remained undetectable, while anti-HCV antibodies were significantly increased over that of day 70.

The results obtained from the sera of chimpanzee 910 were somewhat similar in pattern, although the time of HCV antibodies induced by the infection were not detected during the acute phase of the disease, which extended to at least day 67; the anti-HCV antibodies detected by RIA at day 11 were due to passive immunization of animal 910 with antibodies from the plasma used to inoculate the animal. Anti-HCV antibodies were found in chimpanzee 910 serum during the later, chronic phase of the infection (data not shown).

It should be noted that low ALT values in plasma from individuals with chronic NANBH do not necessarily correlate with weak virus production. A pool of 17 different plasma samples taken from chimpanzee 910 over a period of two to three and one-half years post inoculation was monitored for ALT levels and for HCV RNA. The ALT values of the samples did not exceed 45 mU/ml; nevertheless, titration studies indicated high titers of HCV (3 x $10^6$ CID/ml). cPCR was carried out for 30 cycles, and the autoradiogram was exposed for 15 hours; the cPCR analysis clearly showed the presence of viral RNA (data not shown).

HCV/cPCR Assay of Sera: Detection of Viral RNA in Acut Phase Inf ction in Humans

Plasma from a human surgical patient collected during early acute NANBH was examined for HCV RNA and for anti-HCV antibodies, utilizing the HCV/cPCR assay and C100-RIA, respectively. The HCV/cPCR assay was conducted utilizing 1 microliter of plasma from the patient, and from four human controls with known pedigrees; cPCR was performed for thirty cycles, and after hybridization and washing the autoradiogram was exposed for eight hours.

The results showed that the serum collected from the surgical patient during the acute phase of infection contained a high level of viral RNA, and that anti-HCV antibodies were not detectable by the C100-RIA (data now shown). (The acute phase plasma from the surgical patient was known to have a high titer of NANBH infectious agent [$10^{6.5}$ CID/ml, as determined by Feinstone et al. (1981); Feinstone et al. (1983)]). It should be noted, however, that this patient did sero-covert to anti-HCV antibodies by the C100-RIA approximately 9 months after infection. The serum from the pedigreed human control plasmas were negative in both the HCV/cPCR assay and C100-RIA.

Sensitivity of HCV/cPCR Assay

The sensitivity of the HCV/cPCR assay was determined by analyzing ten-fold serial dilutions of a plasma pool of known titer. The chimpanzee plasma had a titer of ~3 x $10^5$ CID/ml, and RNA was extracted from ten-fold dilutions of 1 microliter of the plasma. cPCR was performed for 30 cycles, and after hybridization and washing, the autoradiogram was exposed for 15 hours. The cPCR products resulting from amplification of ~300, ~30, and ~3 CID of HCV genomes are shown in lanes 1-3, respectively of Fig. 29. The samples in lanes 1 and 2 were detectable on autoradiograms exposed for 2 hours.

Sinc th average titer of HCV in infected individuals is believed to be between approximat ly 100 to 10,000 CID/ml of plasma, this data suggests that the HCV/cPCR assay may be clinically useful.

HCV/cPCR Assay for Variant HCV Strains

Primers, consisting of a set of oligomer 44-mers and a set of oligomer 45-mers, were designed to amplify strains of HCV which are similar or identical to the HCV isolate from which the cDNA sequence in Fig. 18 is derived. The premise underlying the design of these primers is our discovery that HCV is a Flavi-like virus. Members of the Flaviviridae family, when compared to HCV, have two major conserved sets of amino acid sequences, TATPPG and QRRGR, in the putative NS3 region of these viruses. Several other smaller sets may be seen, for example, GDD in the putative NS5 region. Other sets are determinable by comparison of the known amino acid sequences with that of HCV. This information was deduc d from the sequences for several members of Flaviviridae which have been described, including Japanese Encephalitis Virus (Sumiyoshi et al. (1987)), Yellow Fever Virus (Rice et al. (1985)), Dengue Type 2 Virus (Hahn et al. (1988)), Dengue Type 4 Virus (Mackow (1987)), and West Nile Virus (Castle et al. (1986)). The conserved amino acid sequences and codon utilization are in the table immediately following.

C nserved Amino Acid (A.A.) Sequences Among Flaviviruses and HCV

| Virus | # of first A.A. | | T | A | T | P | P | G | |
|---|---|---|---|---|---|---|---|---|---|
| HCV | 1348 | 5' | ACC | GCC | ACC | CCT | CCG | GCC | 3' |
| Yellow Fever | 1805 | | ACA | GCC | ACA | CCG | CCT | GGG | |
| West Nile | 1818 | | ACG | GCA | ACG | CCA | CCC | GGG | |
| Dengue-4 | 1788 | | ACC | GCA | ACC | CCT | CCC | GGA | |
| JEV | 1957 | | ACA | GCG | ACC | CCG | CCT | GGA | |

HCV sense primer (44mer)=
5' ACC GCC ACC CCX CC 3'
(X = A,T,C, or G)

| Virus | # of first A.A. | | Q | R | R | G | R | |
|---|---|---|---|---|---|---|---|---|
| HCV | 1486 | 5' | CAA | CGT | CGG | GGC | AGG | 3' |
| Yellow Fever | 1946 | | CAA | AGG | AGG | GGG | CGC | |
| West Nile | 1959 | | CAG | CGG | AGA | GGA | CGC | |
| Dengue-4 | 1929 | | CAG | AGA | AGA | GGG | CGA | |
| JEV | 1820 | | CAA | CGG | AGG | GGC | AGA | |

HCV antisense primer (45mer)=
3' GTX GCA GCC CCG TCC 5'
(X = T or C)

Note: the primer sequence was chosen to minimize the number of nucleotide degeneracies at the 3'-end of the primer sequence and to maximize the number of nucleotides at the 3'-end of each primer which exactly match any of the possible nucleotide sequences, or the complement thereof, encoding the conserved amino acids indicated above.

The 44-mer and 45-mer oligomer primers were designed so that the sequences encoding these amino acids were incorporated within the primer. Moreover, they contain degeneracies at the 3'-end of each primer, and are derived from two different regions of the HCV genome which are present in clone 40b (See Fig. 28), and which are derived from the region encoding putative NS3 of HCV. The formulae for the oligonucleotide primers in the sets are:

5′ GAC TGC GGG GGC GAG ACT GGT TGT GCT CGC
ACC GCC ACC CCX CC 3′

where X is A,T,G, or C; and

5′ TCT GTA GAT GCC TGG CTT CCC CCT GCC AGT
CCT GCC CCG ACT YTG 3′

where Y is T or C.

The HCV/cPCR assay was carried out utilizing these primers to amplify HCV RNA in chimpanzee 910 plasma. The assay method was essentially as described in Section supra., except that the 44-mer and 45-mer sets of oligom r primers were substituted for the primers derived from clone 36 and clone 37b. In addition, detection of amplified HCV cDNA was by hybridization with a probe derived from clone 40a, the sequence of which is shown in Fig. 32.

The probe was prepared by amplifying a segment of clone 40a utilizing the PCR method described supra., and 18-mer primers containing the following sequences:

5′ GAG ACA TCT CAT CTT CTG 3′

and

5′ GAG CGT GAT TGT CTC AAT 3′.

After amplification, the probe preparation was labeled with $^{32}P$ by nick translation.

Fig. 33 shows an autoradiograph of the Southern blots probed with the sequence derived from Clone 40a. $^{32}P$ labeled MspI digested pBR322 DNA fragments served as markers (lane 1). The predicted size of the PCR product resulting from amplification using these primers is 490

nucle tid s (nt). Duplicate reactions are shown in lan s 2 and 3.

Analysis for Variants of the 5'-Region of HCV

Based upon the Flavivirus model, the 5'-region HCV cDNA which is flanked by the regions represented in clones ag30a and k9-1 encodes a segment of putative envelope and/or matrix protein(s) (E/M). Serum obtained from the chimpanzee from which the HCV cDNA "c" library, was constructed was analyzed by HCV/cPCR to determine whether variants within this target region were present.

The HCV/cPCR assay was performed essentially as described supra., for the isolation of clone 5'-32, except for the primers and probes used. Fig. 37 shows the relationship of the primers and probes (and the clones from which they were derived) to that of the target region of HCV cDNA. One set of PCR primers, ag30a16A and K91Env16B, were derived from clones ag30a and k9-1, which are upstream and downstream, respectively, of the target sequence. The expected size of the cPCR product primed by ag30a16A and K91Env16B is 1.145 kb based upon the confirmed sequence of HCV cDNA. Two other sets of PCR primers covering the region amplified using ag30a16A and K91Env16B, and overlapping each other were also used for PCR amplification of HCV RNA in the serum. Thus, in this case the PCR reactions were run using as one set of primers ag30a16A and CA156e16B, and as the second set of primers CA156e16A and k91Env16B. The expected PCR product sizes for these pairs were 615 nucleotides (NT) and 683 NT, respectively. The table immediately following lists the primer, the clone from which it was derived, and the primer sequence.

Table

| Primer | Clone | Sequenc |
|---|---|---|
| ag30a16A | ag30a | 5′ CTC TAT GGC AAT GAG G 3′ |
| K91Env16B | k9-1 | 5′ CGT TGG CAT AAC TGA T 3′ |
| CA156e16B | 156 | 5′ CGA CAA GAA AGA CAG A 3′ |
| CA156e16A | 156 | 5′ AGC TTC GAC GTC ACA T 3′ |
| CA216a16A | 216 | 5′ TGA ACT ATG CAA CAG G 3′ |
| CA216a16B | 216 | 5′ GGA GTG TGC AGG ATG G 3′ |
| CA84a16A | 84 | 5′ AAG GTT GCA ATT GCT C 3′ |
| CA84a16B | 84 | 5′ ACT AAC AGG ACC TTC G 3′ |

The probes for all of the HCV/cPCR products consisted of $^{32}P$ labeled sections of HCV cDNA which had been prepared by PCR amplification of a region of clone 216 (using CA216a16A and 216a16B as primers), and of clone 84 (using CA84a16A and CA84a16B as primers); $^{32}P$ was introduced into the PCR products by nick translation. These probes did not overlap the primers used in the HCV/cPCR reactions.

Fig. 38 shows an autoradiograph of a Southern blot in which the HCV/cPCR products were hybridized with the $^{32}P$-labeled probes. The HCV/cPCR product extended from primers ag30a16A and K91Env16B (lane 1) was approximately 1.1Kb; no other PCR products were observed in a 15 hour exposure. The HCV products extended from the primer sets ag30a15A/CA156e16B (lane 2) and CA156e16A/K91Env16B (lane 3) were approximately 625NT and approximately 700 NT, respectively. The size of the PCR products were determined by comparison with the relative migrations of fragments resulting from the digestion of pBR322 with MspI and of PhiX 174 digested with HaeIII (lane 5).

The above study will detect insertions or deletions as small as approximately 20NT to 50NT and DNA rearrangements altering the size of the target DNA. The

results in Fig. 38 confirm that there is only 1 major species f cDNA d rived from the E/M r gion of the HCV in the chimpanze serum.

Amplification for Cloning of HCV cDNA Sequences Utilizing the PCR and Primers Derived from Conserved Regions of Flavivirus Genomic Sequences

Our discovery that HCV is a flavi-like virus, allows a strategy for cloning uncharacterized HCV cDNA sequences utilizing the PCR technique, and primers derived from the regions encoding conserved amino acid sequences in flaviviruses. Generally, one of the primers is derived from a defined HCV genomic sequence, and the other primer which flanks a region of unsequenced HCV polynucleotide is derived from a conserved region of the flavivirus genome. The flavivirus genomes are known to contain conserved sequences within the NS1, and E polypeptides, which are encoded in the 5'-region of the flavivirus genome. Thus, to isolate cDNA sequences derived from putatively comparable regions of the HCV genome, upstream primers are designed which are derived from the conserved sequences within these flavivirus polypeptides. The downstream primers are derived from an upstream end of the known portion of the HCV cDNA.

Because of the degeneracy of the code, it is probable that there will be mismatches between the flavivirus probes and the corresponding HCV genomic sequence. Therefore a strategy which is similar to th one described by Lee (1988) is used. The Lee procedure utilizes mixed oligonucleotide primers complementary to the reverse translation products of an amino acid sequence; the sequences in the mixed primers takes into account every codon degeneracy for the conserved amino acid sequence.

Three sets of primer mixes are generated, based on the amino acid homologies found in several

flavivirus s, including D ngu -2,4 (D-2,4), Japanese Encephalitis Virus (JEV), Yellow Fev r (YF), and W st Nile Virus (WN). Th primer mixture d rived from the most upstream conserved sequence (5'-1), is based upon the amino acid sequence gly-trp-gly, which is part of the conserved sequence asp-arg-gly-trp-gly-aspN found in the E protein of D-2, JEV, YF, and WN. The next primer mixture (5'-2) is based upon a downstream conserved sequence in E protein, phe-asp-gly-asp-ser-tyr-ileu-phe-gly-asp-ser-tyr-ileu, and is derived from phe-gly-asp; the conserved sequence is present in D-2, JEV, YF, and WN. The third primer mixture (5'-3), is based on the amino acid sequence arg-ser-cys, which is part of the conserved sequence cys-cys-arg-ser-cys in the NS1 protein of D-2, D-4, JEV, YF, and WN. The individual primers which form the mixture in 5'-3 are shown in Fig. 53. In addition to the varied sequences derived from conserved region, each primer in each mixture also contains a constant region at the 5'-end which contains a sequence encoding sites for restriction enzymes, HindIII, MboI, and EcoRI.

The downstream primer, ssc5h20A, is derived from a nucleotide sequence in clone 5h, which contains HCV cDNA with sequences with overlap those in clones 14i and 11b. The sequence of ssc5h20A is

5' GTA ATA TGG TGA CAG AGT CA 3'.

An alternative primer, ssc5h34A, may also be used. This primer is derived from a sequence in clone 5h, and in addition contains nucleotides at the 5'-end which create a restriction enzyme site, thus facilitating cloning. The sequence of ssc5h34A is

5' GAT CTC TAG AGA AAT CAA TAT GGT GAC AGA GTC A 3'.

The PCR reaction, which was initially describ d by Saiki et al. (1986), is carried out essentially as described in Lee et al. (1988), except that the template for the cDNA is RNA isolated from HCV infected chimpanzee liver, or from viral particles isolated from HCV infected chimpanzee serum. In addition, the annealing conditions are less stringent in the first round of amplification (0.6M NaCl, and 25°C), since the part of the primer which will anneal to the HCV sequence is only 9 nucleotides, and there could be mismatches. Moreover, if ssc5h34A is used, the additional sequences not derived from the HCV genome tend to destabilize the primer-template hybrid. After the first round of amplification, the annealing conditions can be more stringent (0.066M NaCl, and 32°C-37°C), since the amplified sequences now contain regions which are complementary to, or duplicates of the primers. In addition, the first 10 cycles of amplification are run with Klenow enzyme I, under appropriate PCR conditions for that enzyme. After the completion of these cycles, the samples are extracted, and run with Taq polymerase, according to kit directions, as furnished by Cetus/Perkin-Elmer.

After the amplification, the amplified HCV cDNA sequences are detected by hybridization using a probe derived from clone 5h. This probe is derived from sequences upstream of those used to derive the primer, and does not overlap the sequences of the clone 5h derived primers. The sequence of the probe is

5′ CCC AGC GGC GTA CGC GCT GGA CAC GGA GGT GGC CGC GTC
GTG TGG CGG TGT TGT TCT CGT CGG GTT GAT GGC GC 3′.

## Industrial Applicability

The methods described herein, as well as the olig mers, both probes and primers, derived from HCV cDNA, and kits containing them, are useful for the accurate, relatively simple, and economic determination of the presence of HCV in biological samples, more particularly in blood which may be used for transfusions, and in individuals suspected of having HCV an infection. Moreover, these methods and oligomers may be useful for detecting an earlier stage of HCV infection than are immunological assays based upon the use of a recombinant HCV polypeptides. Also, an amplified polynucleotide hybridization assay detects HCV RNA in occasional samples which are anti-HCV antibody negative. Thus, the probes and primers described herein may be used amplified hybridization assays, in conjunction with an immunoassays based on HCV polypeptides to more completely identify infections due to HCV, and HCV-infected biological specimens, including blood.

The information provided herein allows the design of primers and/or probes which are derived from conserved regions of the HCV genome. The provision of these primers and probes makes available a general method which will detect variant HCV strains, and which will be of use in the screening of blood and blood products.

If the primers used in the method are derived from conserved regions of the HCV genome, the method should aid in the detection and/or identification of variant strains of HCV. This, in turn, should lead to the development of additional immunological reagents for the detection and diagnosis of HCV, as well as the developm nt of additional polynucleotide reagents for detection and or treatment of HCV.

In addition, sets of primers and probes designed from the conserved amino acid sequences of Flaviviruses

and HCV allow for a universal d tection method for these inf ctious ag nts.

The following listed materials are on deposit under the terms of the Budapest Treaty with the American Type Culture Collection (ATCC), 12301 Parklawn Dr., Rockville, Maryland 20852, and have been assigned the following Accession Numbers.

| lambda-gt11 | ATCC No. | Deposit Date |
|---|---|---|
| HCV cDNA library | 40394 | 1 Dec. 1987 |
| clone 81 | 40388 | 17 Nov. 1987 |
| clone 91 | 40389 | 17 Nov. 1987 |
| clone 1-2 | 40390 | 17 Nov. 1987 |
| clone 5-1-1 | 40391 | 18 Nov. 1987 |
| clone 12f | 40514 | 10 Nov. 1988 |
| clone 35f | 40511 | 10 Nov. 1988 |
| clone 15e | 40513 | 10 Nov. 1988 |
| clone K9-1 | 40512 | 10 Nov. 1988 |
| JSC 308 | 20879 | 5 May 1988 |
| pS356 | 67683 | 29 April 1988 |

In addition, the following deposits were made on 11 May 1989.

| Strain | Linkers | ATCC No. |
|---|---|---|
| D1210 (Cf1/5-1-1) | EF | 67967 |
| D1210 (Cf1/81) | EF | 67968 |
| D1210 (Cf1/CA74a) | EF | 67969 |
| D1210 (Cf1/35f) | AB | 67970 |
| D1210 (Cf1/279a) | EF | 67971 |
| D1210 (Cf1/C36) | CD | 67972 |
| D1210 (Cf1/13i) | AB | 67973 |
| D1210 (Cf1/C33b) | EF | 67974 |
| D1210 (Cf1/CA290a) | AB | 67975 |
| HB101 (AB24/C100 #3R) | | 67976 |

Th following derivatives of strain D1210 wer d posited on 3 May 1989.

| Strain Derivative | ATCC No. |
|---|---|
| pCF1CS/C8f | 67956 |
| pCF1AB/C12f | 67952 |
| pCF1EF/14c | 67949 |
| pCF1EF/15e | 67954 |
| pCF1AB/C25c | 67958 |
| pCF1EF/C33c | 67953 |
| pCF1EF/C33f | 67050 |
| pCF1CD/33g | 67951 |
| pCF1CD/C39c | 67955 |
| pCF1EF/C40b | 67957 |
| pCF1EF/CA167b | 67959 |

The following strains were deposited on May 12, 1989.

| Strain | ATCC No. |
|---|---|
| Lambda gt11(C35) | 40603 |
| Lambda gt10(beta-5a) | 40602 |
| D1210 (C40b) | 67980 |
| D1210 (M16) | 67981 |

The following biological materials were deposited on March 23, 1990.

| Material | ATCC No. |
|---|---|
| 5'-clone32 (in pUC18S) | 68276 |

## CLAIMS

1. An oligomer capabl of hybridizing to an HCV sequence in an analyte polynucleotide strand, wherein the oligomer is comprised of an HCV targeting sequence complementary to at least 4 contiguous nucleotides of HCV cDNA shown in Fig. 18.

2. The oligomer of claim 1, wherein the targeting sequence is comprised of nucleotides which are complementary to nucleotides selected from the following HCV cDNA nucleotides shown in Fig. 18, ($nn_x$ - $nn_y$ denotes nucleotide number x to nucleotide number y)):

$nn_{-340}$ - $nn_{-330}$; $nn_{-330}$ - $nn_{-320}$; $nn_{-320}$ - $nn_{-310}$;
$nn_{-310}$ - $nn_{-300}$; $nn_{-300}$ - $nn_{-290}$; $nn_{-290}$ - $nn_{-280}$;
$nn_{-280}$ - $nn_{-270}$; $nn_{-270}$ - $nn_{-260}$; $nn_{-260}$ - $nn_{-250}$;
$nn_{-250}$ - $nn_{-240}$; $nn_{-240}$ - $nn_{-230}$; $nn_{-230}$ - $nn_{-220}$;
$nn_{-220}$ - $nn_{-210}$; $nn_{-210}$ - $nn_{-200}$; $nn_{-200}$ - $nn_{-190}$;
$nn_{-190}$ - $nn_{-180}$; $nn_{-180}$ - $nn_{-170}$; $nn_{-170}$ - $nn_{-160}$;
$nn_{-160}$ - $nn_{-150}$; $nn_{-150}$ - $nn_{-140}$; $nn_{-140}$ - $nn_{-130}$;
$nn_{-130}$ - $nn_{-120}$; $nn_{-120}$ - $nn_{-110}$; $nn_{-110}$ - $nn_{-100}$;
$nn_{-100}$ - $nn_{-90}$; $nn_{-90}$ - $nn_{-80}$; $nn_{-80}$ - $nn_{-70}$;
$nn_{-70}$ - $nn_{-60}$; $nn_{-60}$ - $nn_{-50}$; $nn_{-50}$ - $nn_{-40}$;
$nn_{-40}$ - $nn_{-30}$; $nn_{-30}$ - $nn_{-20}$; $nn_{-20}$ - $nn_{-10}$;
$nn_{-10}$ - $nn_{1}$; $nn_{1}$ - $nn_{10}$; $nn_{10}$ - $nn_{20}$; $nn_{20}$ - $nn_{30}$;
$nn_{30}$ - $nn_{40}$; $nn_{40}$ - $nn_{50}$; $nn_{50}$ - $nn_{60}$; $nn_{60}$ - $nn_{70}$;
$nn_{70}$ - $nn_{80}$; $nn_{80}$ - $nn_{90}$; $nn_{90}$ - $nn_{100}$; $nn_{100}$ - $nn_{110}$;
$nn_{110}$ - $nn_{120}$; $nn_{120}$ - $nn_{130}$; $nn_{130}$ - $nn_{140}$;
$nn_{140}$ - $nn_{150}$; $nn_{150}$ - $nn_{160}$; $nn_{160}$ - $nn_{170}$;
$nn_{170}$ - $nn_{180}$; $nn_{180}$ - $nn_{190}$; $nn_{190}$ - $nn_{200}$;
$nn_{200}$ - $nn_{210}$; $nn_{210}$ - $nn_{220}$; $nn_{220}$ - $nn_{230}$;
$nn_{230}$ - $nn_{240}$; $nn_{240}$ - $nn_{250}$; $nn_{250}$ - $nn_{260}$;
$nn_{260}$ - $nn_{270}$; $nn_{270}$ - $nn_{280}$; $nn_{280}$ - $nn_{290}$;
$nn_{290}$ - $nn_{300}$; $nn_{300}$ - $nn_{310}$; $nn_{310}$ - $nn_{320}$;

$nn_{320} - nn_{330}$; $nn_{330} - nn_{340}$; $nn_{340} - nn_{350}$; $nn_{350} - nn_{360}$; $nn_{360} - nn_{370}$; $nn_{370} - nn_{380}$; $nn_{380} - nn_{390}$; $nn_{390} - nn_{400}$; $nn_{400} - nn_{410}$; $nn_{410} - nn_{420}$; $nn_{420} - nn_{430}$; $nn_{430} - nn_{440}$; $nn_{440} - nn_{450}$; $nn_{450} - nn_{460}$; $nn_{460} - nn_{470}$; $nn_{470} - nn_{480}$; $nn_{480} - nn_{490}$; $nn_{490} - nn_{500}$; $nn_{500} - nn_{510}$; $nn_{510} - nn_{520}$; $nn_{520} - nn_{530}$; $nn_{530} - nn_{540}$; $nn_{540} - nn_{550}$; $nn_{550} - nn_{560}$; $nn_{560} - nn_{570}$; $nn_{570} - nn_{580}$; $nn_{580} - nn_{590}$; $nn_{590} - nn_{600}$; $nn_{600} - nn_{610}$; $nn_{610} - nn_{620}$; $nn_{620} - nn_{630}$; $nn_{630} - nn_{640}$; $nn_{640} - nn_{650}$; $nn_{650} - nn_{660}$; $nn_{660} - nn_{670}$; $nn_{670} - nn_{680}$; $nn_{680} - nn_{690}$; $nn_{690} - nn_{700}$; $nn_{700} - nn_{710}$; $nn_{710} - nn_{720}$; $nn_{720} - nn_{730}$; $nn_{730} - nn_{740}$; $nn_{740} - nn_{750}$; $nn_{750} - nn_{760}$; $nn_{760} - nn_{770}$; $nn_{770} - nn_{780}$; $nn_{780} - nn_{790}$; $nn_{790} - nn_{800}$; $nn_{800} - nn_{810}$; $nn_{810} - nn_{820}$; $nn_{820} - nn_{830}$; $nn_{830} - nn_{840}$; $nn_{840} - nn_{850}$; $nn_{850} - nn_{860}$; $nn_{860} - nn_{870}$; $nn_{870} - nn_{880}$; $nn_{880} - nn_{890}$; $nn_{890} - nn_{900}$; $nn_{900} - nn_{910}$; $nn_{910} - nn_{920}$; $nn_{920} - nn_{930}$; $nn_{930} - nn_{940}$; $nn_{940} - nn_{950}$; $nn_{950} - nn_{960}$; $nn_{960} - nn_{970}$; $nn_{970} - nn_{980}$; $nn_{980} - nn_{990}$; $nn_{990} - nn_{1000}$; $nn_{1000} - nn_{1010}$; $nn_{1010} - nn_{1020}$; $nn_{1020} - nn_{1030}$; $nn_{1030} - nn_{1040}$; $nn_{1040} - nn_{1050}$; $nn_{1050} - nn_{1060}$; $nn_{1060} - nn_{1070}$; $nn_{1070} - nn_{1080}$; $nn_{1080} - nn_{1090}$; $nn_{1090} - nn_{1100}$; $nn_{1100} - nn_{1110}$; $nn_{1110} - nn_{1120}$; $nn_{1120} - nn_{1130}$; $nn_{1130} - nn_{1140}$; $nn_{1140} - nn_{1150}$; $nn_{1150} - nn_{1160}$; $nn_{1160} - nn_{1170}$; $nn_{1170} - nn_{1180}$; $nn_{1180} - nn_{1190}$; $nn_{1190} - nn_{1200}$; $nn_{1200} - nn_{1210}$; $nn_{1210} - nn_{1220}$; $nn_{1220} - nn_{1230}$; $nn_{1230} - nn_{1240}$; $nn_{1240} - nn_{1250}$; $nn_{1250} - nn_{1260}$; $nn_{1260} - nn_{1270}$; $nn_{1270} - nn_{1280}$; $nn_{1280} - nn_{1290}$; $nn_{1290} - nn_{1300}$; $nn_{1300} - nn_{1310}$; $nn_{1310} - nn_{1320}$; $nn_{1320} - nn_{1330}$; $nn_{1330} - nn_{1340}$; $nn_{1340} - nn_{1350}$; $nn_{1350} - nn_{1360}$; $nn_{1360} - nn_{1370}$; $nn_{1370} - nn_{1380}$; $nn_{1380} - nn_{1390}$; $nn_{1390} - nn_{1400}$;

$nn_{1400} - nn_{1410}$; $nn_{1410} - nn_{1420}$; $nn_{1420} - nn_{1430}$; $nn_{1430} - nn_{1440}$; $nn_{1440} - nn_{1450}$; $nn_{1450} - nn_{1460}$; $nn_{1460} - nn_{1470}$; $nn_{1470} - nn_{1480}$; $nn_{1480} - nn_{1490}$; $nn_{1490} - nn_{1500}$; $nn_{1500} - nn_{1510}$; $nn_{1510} - nn_{1520}$; $nn_{1520} - nn_{1530}$; $nn_{1530} - nn_{1540}$; $nn_{1540} - nn_{1550}$; $nn_{1550} - nn_{1560}$; $nn_{1560} - nn_{1570}$; $nn_{1570} - nn_{1580}$; $nn_{1580} - nn_{1590}$; $nn_{1590} - nn_{1600}$; $nn_{1600} - nn_{1610}$; $nn_{1610} - nn_{1620}$; $nn_{1620} - nn_{1630}$; $nn_{1630} - nn_{1640}$; $nn_{1640} - nn_{1650}$; $nn_{1650} - nn_{1660}$; $nn_{1660} - nn_{1670}$; $nn_{1670} - nn_{1680}$; $nn_{1680} - nn_{1690}$; $nn_{1690} - nn_{1700}$; $nn_{1700} - nn_{1710}$; $nn_{1710} - nn_{1720}$; $nn_{1720} - nn_{1730}$; $nn_{1730} - nn_{1740}$; $nn_{1740} - nn_{1750}$; $nn_{1750} - nn_{1760}$; $nn_{1760} - nn_{1770}$; $nn_{1770} - nn_{1780}$; $nn_{1780} - nn_{1790}$; $nn_{1790} - nn_{1800}$; $nn_{1800} - nn_{1810}$; $nn_{1810} - nn_{1820}$; $nn_{1820} - nn_{1830}$; $nn_{1830} - nn_{1840}$; $nn_{1840} - nn_{1850}$; $nn_{1850} - nn_{1860}$; $nn_{1860} - nn_{1870}$; $nn_{1870} - nn_{1880}$; $nn_{1880} - nn_{1890}$; $nn_{1890} - nn_{1900}$; $nn_{1900} - nn_{1910}$; $nn_{1910} - nn_{1920}$; $nn_{1920} - nn_{1930}$; $nn_{1930} - nn_{1940}$; $nn_{1940} - nn_{1950}$; $nn_{1950} - nn_{1960}$; $nn_{1960} - nn_{1970}$; $nn_{1970} - nn_{1980}$; $nn_{1980} - nn_{1990}$; $nn_{1990} - nn_{2000}$; $nn_{2000} - nn_{2010}$; $nn_{2010} - nn_{2020}$; $nn_{2020} - nn_{2030}$; $nn_{2030} - nn_{2040}$; $nn_{2040} - nn_{2050}$; $nn_{2050} - nn_{2060}$; $nn_{2060} - nn_{2070}$; $nn_{2070} - nn_{2080}$; $nn_{2080} - nn_{2090}$; $nn_{2090} - nn_{2100}$; $nn_{2100} - nn_{2110}$; $nn_{2110} - nn_{2120}$; $nn_{2120} - nn_{2130}$; $nn_{2130} - nn_{2140}$; $nn_{2140} - nn_{2150}$; $nn_{2150} - nn_{2160}$; $nn_{2160} - nn_{2170}$; $nn_{2170} - nn_{2180}$; $nn_{2180} - nn_{2190}$; $nn_{2190} - nn_{2200}$; $nn_{2200} - nn_{2210}$; $nn_{2210} - nn_{2220}$; $nn_{2220} - nn_{2230}$; $nn_{2230} - nn_{2240}$; $nn_{2240} - nn_{2250}$; $nn_{2250} - nn_{2260}$; $nn_{2260} - nn_{2270}$; $nn_{2270} - nn_{2280}$; $nn_{2280} - nn_{2290}$; $nn_{2290} - nn_{2300}$; $nn_{2300} - nn_{2310}$; $nn_{2310} - nn_{2320}$; $nn_{2320} - nn_{2330}$; $nn_{2330} - nn_{2340}$; $nn_{2340} - nn_{2350}$; $nn_{2350} - nn_{2360}$; $nn_{2360} - nn_{2370}$; $nn_{2370} - nn_{2380}$; $nn_{2380} - nn_{2390}$; $nn_{2390} - nn_{2400}$; $nn_{2400} - nn_{2410}$; $nn_{2410} - nn_{2420}$; $nn_{2420} - nn_{2430}$; $nn_{2430} - nn_{2440}$; $nn_{2440} - nn_{2450}$; $nn_{2450} - nn_{2460}$; $nn_{2460} - nn_{2470}$; $nn_{2470} - nn_{2480}$;

$nn_{2480}$ - $nn_{2490}$; $nn_{2490}$ - $nn_{2500}$; $nn_{2500}$ - $nn_{2510}$; $nn_{2510}$ - $nn_{2520}$; $nn_{2520}$ - $nn_{2530}$; $nn_{2530}$ - $nn_{2540}$; $nn_{2540}$ - $nn_{2550}$; $nn_{2550}$ - $nn_{2560}$; $nn_{2560}$ - $nn_{2570}$; $nn_{2570}$ - $nn_{2580}$; $nn_{2580}$ - $nn_{2590}$; $nn_{2590}$ - $nn_{2600}$; $nn_{2600}$ - $nn_{2610}$; $nn_{2610}$ - $nn_{2620}$; $nn_{2620}$ - $nn_{2630}$; $nn_{2630}$ - $nn_{2640}$; $nn_{2640}$ - $nn_{2650}$; $nn_{2650}$ - $nn_{2660}$; $nn_{2660}$ - $nn_{2670}$; $nn_{2670}$ - $nn_{2680}$; $nn_{2680}$ - $nn_{2690}$; $nn_{2690}$ - $nn_{2700}$; $nn_{2700}$ - $nn_{2710}$; $nn_{2710}$ - $nn_{2720}$; $nn_{2720}$ - $nn_{2730}$; $nn_{2730}$ - $nn_{2740}$; $nn_{2740}$ - $nn_{2750}$; $nn_{2750}$ - $nn_{2760}$; $nn_{2760}$ - $nn_{2770}$; $nn_{2770}$ - $nn_{2780}$; $nn_{2780}$ - $nn_{2790}$; $nn_{2790}$ - $nn_{2800}$; $nn_{2800}$ - $nn_{2810}$; $nn_{2810}$ - $nn_{2820}$; $nn_{2820}$ - $nn_{2830}$; $nn_{2830}$ - $nn_{2840}$; $nn_{2840}$ - $nn_{2850}$; $nn_{2850}$ - $nn_{2860}$; $nn_{2860}$ - $nn_{2870}$; $nn_{2870}$ - $nn_{2880}$; $nn_{2880}$ - $nn_{2890}$; $nn_{2890}$ - $nn_{2900}$; $nn_{2900}$ - $nn_{2910}$; $nn_{2910}$ - $nn_{2920}$; $nn_{2920}$ - $nn_{2930}$; $nn_{2930}$ - $nn_{2940}$; $nn_{2940}$ - $nn_{2950}$; $nn_{2950}$ - $nn_{2960}$; $nn_{2960}$ - $nn_{2970}$; $nn_{2970}$ - $nn_{2980}$; $nn_{2980}$ - $nn_{2990}$; $nn_{2990}$ - $nn_{3000}$; $nn_{3000}$ - $nn_{3010}$; $nn_{3010}$ - $nn_{3020}$; $nn_{3020}$ - $nn_{3030}$; $nn_{3030}$ - $nn_{3040}$; $nn_{3040}$ - $nn_{3050}$; $nn_{3050}$ - $nn_{3060}$; $nn_{3060}$ - $nn_{3070}$; $nn_{3070}$ - $nn_{3080}$; $nn_{3080}$ - $nn_{3090}$; $nn_{3090}$ - $nn_{3100}$; $nn_{3100}$ - $nn_{3110}$; $nn_{3110}$ - $nn_{3120}$; $nn_{3120}$ - $nn_{3130}$; $nn_{3130}$ - $nn_{3140}$; $nn_{3140}$ - $nn_{3150}$; $nn_{3150}$ - $nn_{3160}$; $nn_{3160}$ - $nn_{3170}$; $nn_{3170}$ - $nn_{3180}$; $nn_{3180}$ - $nn_{3190}$; $nn_{3190}$ - $nn_{3200}$; $nn_{3200}$ - $nn_{3210}$; $nn_{3210}$ - $nn_{3220}$; $nn_{3220}$ - $nn_{3230}$; $nn_{3230}$ - $nn_{3240}$; $nn_{3240}$ - $nn_{3250}$; $nn_{3250}$ - $nn_{3260}$; $nn_{3260}$ - $nn_{3270}$; $nn_{3270}$ - $nn_{3280}$; $nn_{3280}$ - $nn_{3290}$; $nn_{3290}$ - $nn_{3300}$; $nn_{3300}$ - $nn_{3310}$; $nn_{3310}$ - $nn_{3320}$; $nn_{3320}$ - $nn_{3330}$; $nn_{3330}$ - $nn_{3340}$; $nn_{3340}$ - $nn_{3350}$; $nn_{3350}$ - $nn_{3360}$; $nn_{3360}$ - $nn_{3370}$; $nn_{3370}$ - $nn_{3380}$; $nn_{3380}$ - $nn_{3390}$; $nn_{3390}$ - $nn_{3400}$; $nn_{3400}$ - $nn_{3410}$; $nn_{3410}$ - $nn_{3420}$; $nn_{3420}$ - $nn_{3430}$; $nn_{3430}$ - $nn_{3440}$; $nn_{3440}$ - $nn_{3450}$; $nn_{3450}$ - $nn_{3460}$; $nn_{3460}$ - $nn_{3470}$; $nn_{3470}$ - $nn_{3480}$; $nn_{3480}$ - $nn_{3490}$; $nn_{3490}$ - $nn_{3500}$; $nn_{3500}$ - $nn_{3510}$; $nn_{3510}$ - $nn_{3520}$; $nn_{3520}$ - $nn_{3530}$; $nn_{3530}$ - $nn_{3540}$; $nn_{3540}$ - $nn_{3550}$; $nn_{3550}$ - $nn_{3560}$;

$nn_{3560} - nn_{3570}$; $nn_{3570} - nn_{3580}$; $nn_{3580} - nn_{3590}$;
$nn_{3590} - nn_{3600}$; $nn_{3600} - nn_{3610}$; $nn_{3610} - nn_{3620}$;
$nn_{3620} - nn_{3630}$; $nn_{3630} - nn_{3640}$; $nn_{3640} - nn_{3650}$;
$nn_{3650} - nn_{3660}$; $nn_{3660} - nn_{3670}$; $nn_{3670} - nn_{3680}$;
$nn_{3680} - nn_{3690}$; $nn_{3690} - nn_{3700}$; $nn_{3700} - nn_{3710}$;
$nn_{3710} - nn_{3720}$; $nn_{3720} - nn_{3730}$; $nn_{3730} - nn_{3740}$;
$nn_{3740} - nn_{3750}$; $nn_{3750} - nn_{3760}$; $nn_{3760} - nn_{3770}$;
$nn_{3770} - nn_{3780}$; $nn_{3780} - nn_{3790}$; $nn_{3790} - nn_{3800}$;
$nn_{3800} - nn_{3810}$; $nn_{3810} - nn_{3820}$; $nn_{3820} - nn_{3830}$;
$nn_{3830} - nn_{3840}$; $nn_{3840} - nn_{3850}$; $nn_{3850} - nn_{3860}$;
$nn_{3860} - nn_{3870}$; $nn_{3870} - nn_{3880}$; $nn_{3880} - nn_{3890}$;
$nn_{3890} - nn_{3900}$; $nn_{3900} - nn_{3910}$; $nn_{3910} - nn_{3920}$;
$nn_{3920} - nn_{3930}$; $nn_{3930} - nn_{3940}$; $nn_{3940} - nn_{3950}$;
$nn_{3950} - nn_{3960}$; $nn_{3960} - nn_{3970}$; $nn_{3970} - nn_{3980}$;
$nn_{3980} - nn_{3990}$; $nn_{3990} - nn_{4000}$; $nn_{4000} - nn_{4010}$;
$nn_{4010} - nn_{4020}$; $nn_{4020} - nn_{4030}$; $nn_{4030} - nn_{4040}$;
$nn_{4040} - nn_{4050}$; $nn_{4050} - nn_{4060}$; $nn_{4060} - nn_{4070}$;
$nn_{4070} - nn_{4080}$; $nn_{4080} - nn_{4090}$; $nn_{4090} - nn_{4100}$;
$nn_{4100} - nn_{4110}$; $nn_{4110} - nn_{4120}$; $nn_{4120} - nn_{4130}$;
$nn_{4130} - nn_{4140}$; $nn_{4140} - nn_{4150}$; $nn_{4150} - nn_{4160}$;
$nn_{4160} - nn_{4170}$; $nn_{4170} - nn_{4180}$; $nn_{4180} - nn_{4190}$;
$nn_{4190} - nn_{4200}$; $nn_{4200} - nn_{4210}$; $nn_{4210} - nn_{4220}$;
$nn_{4220} - nn_{4230}$; $nn_{4230} - nn_{4240}$; $nn_{4240} - nn_{4250}$;
$nn_{4250} - nn_{4260}$; $nn_{4260} - nn_{4270}$; $nn_{4270} - nn_{4280}$;
$nn_{4280} - nn_{4290}$; $nn_{4290} - nn_{4300}$; $nn_{4300} - nn_{4310}$;
$nn_{4310} - nn_{4320}$; $nn_{4320} - nn_{4330}$; $nn_{4330} - nn_{4340}$;
$nn_{4340} - nn_{4350}$; $nn_{4350} - nn_{4360}$; $nn_{4360} - nn_{4370}$;
$nn_{4370} - nn_{4380}$; $nn_{4380} - nn_{4390}$; $nn_{4390} - nn_{4400}$;
$nn_{4400} - nn_{4410}$; $nn_{4410} - nn_{4420}$; $nn_{4420} - nn_{4430}$;
$nn_{4430} - nn_{4440}$; $nn_{4440} - nn_{4450}$; $nn_{4450} - nn_{4460}$;
$nn_{4460} - nn_{4470}$; $nn_{4470} - nn_{4480}$; $nn_{4480} - nn_{4490}$;
$nn_{4490} - nn_{4500}$; $nn_{4500} - nn_{4510}$; $nn_{4510} - nn_{4520}$;
$nn_{4520} - nn_{4530}$; $nn_{4530} - nn_{4540}$; $nn_{4540} - nn_{4550}$;
$nn_{4550} - nn_{4560}$; $nn_{4560} - nn_{4570}$; $nn_{4570} - nn_{4580}$;
$nn_{4580} - nn_{4590}$; $nn_{4590} - nn_{4600}$; $nn_{4600} - nn_{4610}$;
$nn_{4610} - nn_{4620}$; $nn_{4620} - nn_{4630}$; $nn_{4630} - nn_{4640}$;

$nn_{4640}$ - $nn_{4650}$; $nn_{4650}$ - $nn_{4660}$; $nn_{4660}$ - $nn_{4670}$;
$nn_{4670}$ - $nn_{4680}$; $nn_{4680}$ - $nn_{4690}$; $nn_{4690}$ - $nn_{4700}$;
$nn_{4700}$ - $nn_{4710}$; $nn_{4710}$ - $nn_{4720}$; $nn_{4720}$ - $nn_{4730}$;
$nn_{4730}$ - $nn_{4740}$; $nn_{4740}$ - $nn_{4750}$; $nn_{4750}$ - $nn_{4760}$;
$nn_{4760}$ - $nn_{4770}$; $nn_{4770}$ - $nn_{4780}$; $nn_{4780}$ - $nn_{4790}$;
$nn_{4790}$ - $nn_{4800}$; $nn_{4800}$ - $nn_{4810}$; $nn_{4810}$ - $nn_{4820}$;
$nn_{4820}$ - $nn_{4830}$; $nn_{4830}$ - $nn_{4840}$; $nn_{4840}$ - $nn_{4850}$;
$nn_{4850}$ - $nn_{4860}$; $nn_{4860}$ - $nn_{4870}$; $nn_{4870}$ - $nn_{4880}$;
$nn_{4880}$ - $nn_{4890}$; $nn_{4890}$ - $nn_{4900}$; $nn_{4900}$ - $nn_{4910}$;
$nn_{4910}$ - $nn_{4920}$; $nn_{4920}$ - $nn_{4930}$; $nn_{4930}$ - $nn_{4940}$;
$nn_{4940}$ - $nn_{4950}$; $nn_{4950}$ - $nn_{4960}$; $nn_{4960}$ - $nn_{4970}$;
$nn_{4970}$ - $nn_{4980}$; $nn_{4980}$ - $nn_{4990}$; $nn_{4990}$ - $nn_{5000}$;
$nn_{5000}$ - $nn_{5010}$; $nn_{5010}$ - $nn_{5020}$; $nn_{5020}$ - $nn_{5030}$;
$nn_{5030}$ - $nn_{5040}$; $nn_{5040}$ - $nn_{5050}$; $nn_{5050}$ - $nn_{5060}$;
$nn_{5060}$ - $nn_{5070}$; $nn_{5070}$ - $nn_{5080}$; $nn_{5080}$ - $nn_{5090}$;
$nn_{5090}$ - $nn_{5100}$; $nn_{5100}$ - $nn_{5110}$; $nn_{5110}$ - $nn_{5120}$;
$nn_{5120}$ - $nn_{5130}$; $nn_{5130}$ - $nn_{5140}$; $nn_{5140}$ - $nn_{5150}$;
$nn_{5150}$ - $nn_{5160}$; $nn_{5160}$ - $nn_{5170}$; $nn_{5170}$ - $nn_{5180}$;
$nn_{5180}$ - $nn_{5190}$; $nn_{5190}$ - $nn_{5200}$; $nn_{5200}$ - $nn_{5210}$;
$nn_{5210}$ - $nn_{5220}$; $nn_{5220}$ - $nn_{5230}$; $nn_{5230}$ - $nn_{5240}$;
$nn_{5240}$ - $nn_{5250}$; $nn_{5250}$ - $nn_{5260}$; $nn_{5260}$ - $nn_{5270}$;
$nn_{5270}$ - $nn_{5280}$; $nn_{5280}$ - $nn_{5290}$; $nn_{5290}$ - $nn_{5300}$;
$nn_{5300}$ - $nn_{5310}$; $nn_{5310}$ - $nn_{5320}$; $nn_{5320}$ - $nn_{5330}$;
$nn_{5330}$ - $nn_{5340}$; $nn_{5340}$ - $nn_{5350}$; $nn_{5350}$ - $nn_{5360}$;
$nn_{5360}$ - $nn_{5370}$; $nn_{5370}$ - $nn_{5380}$; $nn_{5380}$ - $nn_{5390}$;
$nn_{5390}$ - $nn_{5400}$; $nn_{5400}$ - $nn_{5410}$; $nn_{5410}$ - $nn_{5420}$;
$nn_{5420}$ - $nn_{5430}$; $nn_{5430}$ - $nn_{5440}$; $nn_{5440}$ - $nn_{5450}$;
$nn_{5450}$ - $nn_{5460}$; $nn_{5460}$ - $nn_{5470}$; $nn_{5470}$ - $nn_{5480}$;
$nn_{5480}$ - $nn_{5490}$; $nn_{5490}$ - $nn_{5500}$; $nn_{5500}$ - $nn_{5510}$;
$nn_{5510}$ - $nn_{5520}$; $nn_{5520}$ - $nn_{5530}$; $nn_{5530}$ - $nn_{5540}$;
$nn_{5540}$ - $nn_{5550}$; $nn_{5550}$ - $nn_{5560}$; $nn_{5560}$ - $nn_{5570}$;
$nn_{5570}$ - $nn_{5580}$; $nn_{5580}$ - $nn_{5590}$; $nn_{5590}$ - $nn_{5600}$;
$nn_{5600}$ - $nn_{5610}$; $nn_{5610}$ - $nn_{5620}$; $nn_{5620}$ - $nn_{5630}$;
$nn_{5630}$ - $nn_{5640}$; $nn_{5640}$ - $nn_{5650}$; $nn_{5650}$ - $nn_{5660}$;
$nn_{5660}$ - $nn_{5670}$; $nn_{5670}$ - $nn_{5680}$; $nn_{5680}$ - $nn_{5690}$;
$nn_{5690}$ - $nn_{5700}$; $nn_{5700}$ - $nn_{5710}$; $nn_{5710}$ - $nn_{5720}$;

$nn_{5720} - nn_{5730}$; $nn_{5730} - nn_{5740}$; $nn_{5740} - nn_{5750}$; $nn_{5750} - nn_{5760}$; $nn_{5760} - nn_{5770}$; $nn_{5770} - nn_{5780}$; $nn_{5780} - nn_{5790}$; $nn_{5790} - nn_{5800}$; $nn_{5800} - nn_{5810}$; $nn_{5810} - nn_{5820}$; $nn_{5820} - nn_{5830}$; $nn_{5830} - nn_{5840}$; $nn_{5840} - nn_{5850}$; $nn_{5850} - nn_{5860}$; $nn_{5860} - nn_{5870}$; $nn_{5870} - nn_{5880}$; $nn_{5880} - nn_{5890}$; $nn_{5890} - nn_{5900}$; $nn_{5900} - nn_{5910}$; $nn_{5910} - nn_{5920}$; $nn_{5920} - nn_{5930}$; $nn_{5930} - nn_{5940}$; $nn_{5940} - nn_{5950}$; $nn_{5950} - nn_{5960}$; $nn_{5960} - nn_{5970}$; $nn_{5970} - nn_{5980}$; $nn_{5980} - nn_{5990}$; $nn_{5990} - nn_{6000}$; $nn_{6000} - nn_{6010}$; $nn_{6010} - nn_{6020}$; $nn_{6020} - nn_{6030}$; $nn_{6030} - nn_{6040}$; $nn_{6040} - nn_{6050}$; $nn_{6050} - nn_{6060}$; $nn_{6060} - nn_{6070}$; $nn_{6070} - nn_{6080}$; $nn_{6080} - nn_{6090}$; $nn_{6090} - nn_{6100}$; $nn_{6100} - nn_{6110}$; $nn_{6110} - nn_{6120}$; $nn_{6120} - nn_{6130}$; $nn_{6130} - nn_{6140}$; $nn_{6140} - nn_{6150}$; $nn_{6150} - nn_{6160}$; $nn_{6160} - nn_{6170}$; $nn_{6170} - nn_{6180}$; $nn_{6180} - nn_{6190}$; $nn_{6190} - nn_{6200}$; $nn_{6200} - nn_{6210}$; $nn_{6210} - nn_{6220}$; $nn_{6220} - nn_{6230}$; $nn_{6230} - nn_{6240}$; $nn_{6240} - nn_{6250}$; $nn_{6250} - nn_{6260}$; $nn_{6260} - nn_{6270}$; $nn_{6270} - nn_{6280}$; $nn_{6280} - nn_{6290}$; $nn_{6290} - nn_{6300}$; $nn_{6300} - nn_{6310}$; $nn_{6310} - nn_{6320}$; $nn_{6320} - nn_{6330}$; $nn_{6330} - nn_{6340}$; $nn_{6340} - nn_{6350}$; $nn_{6350} - nn_{6360}$; $nn_{6360} - nn_{6370}$; $nn_{6370} - nn_{6380}$; $nn_{6380} - nn_{6390}$; $nn_{6390} - nn_{6400}$; $nn_{6400} - nn_{6410}$; $nn_{6410} - nn_{6420}$; $nn_{6420} - nn_{6430}$; $nn_{6430} - nn_{6440}$; $nn_{6440} - nn_{6450}$; $nn_{6450} - nn_{6460}$; $nn_{6460} - nn_{6470}$; $nn_{6470} - nn_{6480}$; $nn_{6480} - nn_{6490}$; $nn_{6490} - nn_{6500}$; $nn_{6500} - nn_{6510}$; $nn_{6510} - nn_{6520}$; $nn_{6520} - nn_{6530}$; $nn_{6530} - nn_{6540}$; $nn_{6540} - nn_{6550}$; $nn_{6550} - nn_{6560}$; $nn_{6560} - nn_{6570}$; $nn_{6570} - nn_{6580}$; $nn_{6580} - nn_{6590}$; $nn_{6590} - nn_{6600}$; $nn_{6600} - nn_{6610}$; $nn_{6610} - nn_{6620}$; $nn_{6620} - nn_{6630}$; $nn_{6630} - nn_{6640}$; $nn_{6640} - nn_{6650}$; $nn_{6650} - nn_{6660}$; $nn_{6660} - nn_{6670}$; $nn_{6670} - nn_{6680}$; $nn_{6680} - nn_{6690}$; $nn_{6690} - nn_{6700}$; $nn_{6700} - nn_{6710}$; $nn_{6710} - nn_{6720}$; $nn_{6720} - nn_{6730}$; $nn_{6730} - nn_{6740}$; $nn_{6740} - nn_{6750}$; $nn_{6750} - nn_{6760}$; $nn_{6760} - nn_{6770}$; $nn_{6770} - nn_{6780}$; $nn_{6780} - nn_{6790}$; $nn_{6790} - nn_{6800}$;

$nn_{6800}$ - $nn_{6810}$; $nn_{6810}$ - $nn_{6820}$; $nn_{6820}$ - $nn_{6830}$;
$nn_{6830}$ - $nn_{6840}$; $nn_{6840}$ - $nn_{6850}$; $nn_{6850}$ - $nn_{6860}$;
$nn_{6860}$ - $nn_{6870}$; $nn_{6870}$ - $nn_{6880}$; $nn_{6880}$ - $nn_{6890}$;
$nn_{6890}$ - $nn_{6900}$; $nn_{6900}$ - $nn_{6910}$; $nn_{6910}$ - $nn_{6920}$;
$nn_{6920}$ - $nn_{6930}$; $nn_{6930}$ - $nn_{6940}$; $nn_{6940}$ - $nn_{6950}$;
$nn_{6950}$ - $nn_{6960}$; $nn_{6960}$ - $nn_{6970}$; $nn_{6970}$ - $nn_{6980}$;
$nn_{6980}$ - $nn_{6990}$; $nn_{6990}$ - $nn_{7000}$; $nn_{7000}$ - $nn_{7010}$;
$nn_{7010}$ - $nn_{7020}$; $nn_{7020}$ - $nn_{7030}$; $nn_{7030}$ - $nn_{7040}$;
$nn_{7040}$ - $nn_{7050}$; $nn_{7050}$ - $nn_{7060}$; $nn_{7060}$ - $nn_{7070}$;
$nn_{7070}$ - $nn_{7080}$; $nn_{7080}$ - $nn_{7090}$; $nn_{7090}$ - $nn_{7100}$;
$nn_{7100}$ - $nn_{7110}$; $nn_{7110}$ - $nn_{7120}$; $nn_{7120}$ - $nn_{7130}$;
$nn_{7130}$ - $nn_{7140}$; $nn_{7140}$ - $nn_{7150}$; $nn_{7150}$ - $nn_{7160}$;
$nn_{7160}$ - $nn_{7170}$; $nn_{7170}$ - $nn_{7180}$; $nn_{7180}$ - $nn_{7190}$;
$nn_{7190}$ - $nn_{7200}$; $nn_{7200}$ - $nn_{7210}$; $nn_{7210}$ - $nn_{7220}$;
$nn_{7220}$ - $nn_{7230}$; $nn_{7230}$ - $nn_{7240}$; $nn_{7240}$ - $nn_{7250}$;
$nn_{7250}$ - $nn_{7260}$; $nn_{7260}$ - $nn_{7270}$; $nn_{7270}$ - $nn_{7280}$;
$nn_{7280}$ - $nn_{7290}$; $nn_{7290}$ - $nn_{7300}$; $nn_{7300}$ - $nn_{7310}$;
$nn_{7310}$ - $nn_{7320}$; $nn_{7320}$ - $nn_{7330}$; $nn_{7330}$ - $nn_{7340}$;
$nn_{7340}$ - $nn_{7350}$; $nn_{7350}$ - $nn_{7360}$; $nn_{7360}$ - $nn_{7370}$;
$nn_{7370}$ - $nn_{7380}$; $nn_{7380}$ - $nn_{7390}$; $nn_{7390}$ - $nn_{7400}$;
$nn_{7400}$ - $nn_{7410}$; $nn_{7410}$ - $nn_{7420}$; $nn_{7420}$ - $nn_{7430}$;
$nn_{7430}$ - $nn_{7440}$; $nn_{7440}$ - $nn_{7450}$; $nn_{7450}$ - $nn_{7460}$;
$nn_{7460}$ - $nn_{7470}$; $nn_{7470}$ - $nn_{7480}$; $nn_{7480}$ - $nn_{7490}$;
$nn_{7490}$ - $nn_{7500}$; $nn_{7500}$ - $nn_{7510}$; $nn_{7510}$ - $nn_{7520}$;
$nn_{7520}$ - $nn_{7530}$; $nn_{7530}$ - $nn_{7540}$; $nn_{7540}$ - $nn_{7550}$;
$nn_{7550}$ - $nn_{7560}$; $nn_{7560}$ - $nn_{7570}$; $nn_{7570}$ - $nn_{7580}$;
$nn_{7580}$ - $nn_{7590}$; $nn_{7590}$ - $nn_{7600}$; $nn_{7600}$ - $nn_{7610}$;
$nn_{7610}$ - $nn_{7620}$; $nn_{7620}$ - $nn_{7630}$; $nn_{7630}$ - $nn_{7640}$;
$nn_{7640}$ - $nn_{7650}$; $nn_{7650}$ - $nn_{7660}$; $nn_{7660}$ - $nn_{7670}$;
$nn_{7670}$ - $nn_{7680}$; $nn_{7680}$ - $nn_{7690}$; $nn_{7690}$ - $nn_{7700}$;
$nn_{7700}$ - $nn_{7710}$; $nn_{7710}$ - $nn_{7720}$; $nn_{7720}$ - $nn_{7730}$;
$nn_{7730}$ - $nn_{7740}$; $nn_{7740}$ - $nn_{7750}$; $nn_{7750}$ - $nn_{7760}$;
$nn_{7760}$ - $nn_{7770}$; $nn_{7770}$ - $nn_{7780}$; $nn_{7780}$ - $nn_{7790}$;
$nn_{7790}$ - $nn_{7800}$; $nn_{7800}$ - $nn_{7810}$; $nn_{7810}$ - $nn_{7820}$;
$nn_{7820}$ - $nn_{7830}$; $nn_{7830}$ - $nn_{7840}$; $nn_{7840}$ - $nn_{7850}$;
$nn_{7850}$ - $nn_{7860}$; $nn_{7860}$ - $nn_{7870}$; $nn_{7870}$ - $nn_{7880}$;

$nn_{7880}$ - $nn_{7890}$; $nn_{7890}$ - $nn_{7900}$; $nn_{7900}$ - $nn_{7910}$; $nn_{7910}$ - $nn_{7920}$; $nn_{7920}$ - $nn_{7930}$; $nn_{7930}$ - $nn_{7940}$; $nn_{7940}$ - $nn_{7950}$; $nn_{7950}$ - $nn_{7960}$; $nn_{7960}$ - $nn_{7970}$; $nn_{7970}$ - $nn_{7980}$; $nn_{7980}$ - $nn_{7990}$; $nn_{7990}$ - $nn_{8000}$; $nn_{8000}$ - $nn_{8010}$; $nn_{8010}$ - $nn_{8020}$; $nn_{8020}$ - $nn_{8030}$; $nn_{8030}$ - $nn_{8040}$; $nn_{8040}$ - $nn_{8050}$; $nn_{8050}$ - $nn_{8060}$; $nn_{8060}$ - $nn_{8070}$; $nn_{8070}$ - $nn_{8080}$; $nn_{8080}$ - $nn_{8090}$; $nn_{8090}$ - $nn_{8100}$; $nn_{8100}$ - $nn_{8110}$; $nn_{8110}$ - $nn_{8120}$; $nn_{8120}$ - $nn_{8130}$; $nn_{8130}$ - $nn_{8140}$; $nn_{8140}$ - $nn_{8150}$; $nn_{8150}$ - $nn_{8160}$; $nn_{8160}$ - $nn_{8170}$; $nn_{8170}$ - $nn_{8180}$; $nn_{8180}$ - $nn_{8190}$; $nn_{8190}$ - $nn_{8200}$; $nn_{8200}$ - $nn_{8210}$; $nn_{8210}$ - $nn_{8220}$; $nn_{8220}$ - $nn_{8230}$; $nn_{8230}$ - $nn_{8240}$; $nn_{8240}$ - $nn_{8250}$; $nn_{8250}$ - $nn_{8260}$; $nn_{8260}$ - $nn_{8270}$; $nn_{8270}$ - $nn_{8280}$; $nn_{8280}$ - $nn_{8290}$; $nn_{8290}$ - $nn_{8300}$; $nn_{8300}$ - $nn_{8310}$; $nn_{8310}$ - $nn_{8320}$; $nn_{8320}$ - $nn_{8330}$; $nn_{8330}$ - $nn_{8340}$; $nn_{8340}$ - $nn_{8350}$; $nn_{8350}$ - $nn_{8360}$; $nn_{8360}$ - $nn_{8370}$; $nn_{8370}$ - $nn_{8380}$; $nn_{8380}$ - $nn_{8390}$; $nn_{8390}$ - $nn_{8400}$; $nn_{8400}$ - $nn_{8410}$; $nn_{8410}$ - $nn_{8420}$; $nn_{8420}$ - $nn_{8430}$; $nn_{8430}$ - $nn_{8440}$; $nn_{8440}$ - $nn_{8450}$; $nn_{8450}$ - $nn_{8460}$; $nn_{8460}$ - $nn_{8470}$; $nn_{8470}$ - $nn_{8480}$; $nn_{8480}$ - $nn_{8490}$; $nn_{8490}$ - $nn_{8500}$; $nn_{8500}$ - $nn_{8510}$; $nn_{8510}$ - $nn_{8520}$; $nn_{8520}$ - $nn_{8530}$; $nn_{8530}$ - $nn_{8540}$; $nn_{8540}$ - $nn_{8550}$; $nn_{8550}$ - $nn_{8560}$; $nn_{8560}$ - $nn_{8570}$; $nn_{8570}$ - $nn_{8580}$; $nn_{8580}$ - $nn_{8590}$; $nn_{8590}$ - $nn_{8600}$; $nn_{8600}$ - $nn_{8610}$; $nn_{8610}$ - $nn_{8620}$; $nn_{8620}$ - $nn_{8630}$; $nn_{8630}$ - $nn_{8640}$; $nn_{8640}$ - $nn_{8650}$; $nn_{8650}$ - $nn_{8660}$; $nn_{8660}$ - $nn_{8670}$; $nn_{8670}$ - $nn_{8680}$; $nn_{8680}$ - $nn_{8690}$; $nn_{8690}$ - $nn_{8700}$; $nn_{8700}$ - $nn_{8710}$; $nn_{8710}$ - $nn_{8720}$; $nn_{8720}$ - $nn_{8730}$; $nn_{8730}$ - $nn_{8740}$; $nn_{8740}$ - $nn_{8750}$; $nn_{8750}$ - $nn_{8760}$; $nn_{8760}$ - $nn_{8770}$; $nn_{8770}$ - $nn_{8780}$; $nn_{8780}$ - $nn_{8790}$; $nn_{8790}$ - $nn_{8800}$; $nn_{8800}$ - $nn_{8810}$; $nn_{8810}$ - $nn_{8820}$; $nn_{8820}$ - $nn_{8830}$; $nn_{8830}$ - $nn_{8840}$; $nn_{8840}$ - $nn_{8850}$; $nn_{8850}$ - $nn_{8860}$; $nn_{8860}$ - $nn_{8870}$; $nn_{8870}$ - $nn_{8880}$; $nn_{8880}$ - $nn_{8890}$; $nn_{8890}$ - $nn_{8900}$; $nn_{8900}$ - $nn_{8910}$; $nn_{8910}$ - $nn_{8920}$; $nn_{8920}$ - $nn_{8930}$; $nn_{8930}$ - $nn_{8940}$; $nn_{8940}$ - $nn_{8950}$; $nn_{8950}$ - $nn_{8960}$;

$nn_{8960}$ - $nn_{8970}$; $nn_{8970}$ - $nn_{8980}$; $nn_{8980}$ - $nn_{8990}$; $nn_{8990}$ - $nn_{9000}$; $nn_{9000}$ - $nn_{9010}$; $nn_{9010}$ - $nn_{9020}$; $nn_{9020}$ - $nn_{9030}$; $nn_{9030}$ - $nn_{9040}$; $nn_{9040}$ - $nn_{9050}$; $nn_{9050}$ - $nn_{9060}$.

3. The oligomer of claim 1, wherein the targeting sequence is comprised of a sequence which is complementary to a sequence of at least 8 nucleotides present in a conserved HCV nucleotide sequence in HCV RNA.

4. The oligomer of claim 3, wherein the conserved sequence is located in the sequence of nucleotide numbers from the 5'-terminus to about 200 in Fig. 18.

5. The oligomer of claim 3, wherein the conserved sequence is located in the sequence of nucleotide numbers from about 4000 to about 5000 in Fig. 18.

6. The oligomer of claim 3, wherein the conserved sequence is located in the sequence of nucleotide numbers from about 8000 to about 9040 as shown in Fig. 18.

7. The oligomer of claim 3, wherein the conserved sequence is located in the sequence of nucleotide numbers from about -318 to about 174 as shown in Fig. 18.

8. The oligomer of claim 3, wherein the conserved sequence is located in the sequence of nucleotide numbers from about or from about 4056 to about 4448 as shown in Fig. 18.

9. The oligomer of claim 3, wh r in th cons rved sequence is located in the sequence of nucle tide numbers from about 4378 to about 4902 as shown in Fig. 18.

10. The oligomer of claim 3, wherein the conserved sequence is located in the sequence of nucleotide numbers from about 4042 to about 4059 as shown in Fig. 18.

11. The oligomer of claim 3, wherein the conserved sequence is located in the sequence of nucleotide numbers from about 4456 to about 4470, as shown in Fig. 18.

12. The oligomer of claim 3, wherein the conserved sequence is located in the sequence of nucleotide numbers from about 8209 to about 8217, as shown in Fig. 18.

13. The oligomer of claim 3, which is a capture probe.

14. The oligomer of claim 3, which is a label probe.

15. The oligomer of claim 3, which is a primer.

16. A process for detecting an HCV sequence in an analyte strand suspected of containing an HCV polynucleotide, wherein the HCV polynucleotide comprises a selected target region, said process comprising:

(a) providing an oligomer capable of hybridizing to an HCV sequence in an analyte polynucleotide strand, wherein the oligomer is comprised of an HCV targeting

s qu nc complementary to at least 4 contiguous nucleotides of HCV cDNA shown in Fig. 18

(b) incubating th analyte strand with th oligomer of (a) which allow specific hybrid duplexes to form between the targeting sequence and the target sequence; and

(d) detecting hybrids formed between target region, if any, and the oligomer.

17. The process of claim 16 which further comprises:

(a) providing a set oligomers which are primers for the polymerase chain reaction method and which flank the target region; and

(b) amplifying the target region via a polymerase chain reaction method.

18. A kit for detecting an HCV target sequenc in an analyte strand, comprising the oligomer of claim 1 packaged in a suitable container.

19. A method for preparing blood free of HCV comprising:

(a) providing analyte nucleic acids from a sample of blood suspected of containing an HCV target sequence;

(b) providing an oligomer capable of hybridizing to the HCV sequence in an analyte polynucleotide strand, if any, wherein the oligomer is comprised of an HCV targeting sequence complementary to a sequence of at least 8 nucleotides present in a conserved HCV nucleotide sequence in HCV RNA;

(c) reacting (a) with (b) under conditions which allow the formation of a polynucleotide duplex

b tw en the targeting sequence and the target sequence, if any;

(d) detecting a duplex formed in (c), if any; and

(e) saving the blood from which complexes were not detected in (d).

FIG. 1 Translation of DNA 12f

```
    IlePheLysIleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsn
  1 CCATATTTAAAATCAGGATGTACGTGGGAGGGGTCGAACACAGGCTGGAAGCTGCCTGCA
    GGTATAAATTTTAGTCCTACATGCACCCTCCCCAGCTTGTGTCCGACCTTCGACGGACGT

    TrpThrArgGlyGluArgCysAspLeuGluAspArgAspArgSerGluLeuSerProLeu
 61 ACTGGACGCGGGGCGAACGTTGCGATCTGGAAGACAGGGACAGGTCCGAGCTCAGCCCGT
    TGACCTGCGCCCCGCTTGCAACGCTAGACCTTCTGTCCCTGTCCAGGCTCGAGTCGGGCA

    LeuLeuThrThrThrGlnTrpGlnValLeuProCysSerPheThrThrLeuProAlaLeu
121 TACTGCTGACCACTACACAGTGGCAGGTCCTCCCGTGTTCCTTCACAACCCTACCAGCCT
    ATGACGACTGGTGATGTGTCACCGTCCAGGAGGGCACAAGGAAGTGTTGGGATGGTCGGA

    SerThrGlyLeuIleHisLeuHisGlnAsnIleValAspValGlnTyrLeuTyrGlyVal
181 TGTCCACCGGCCTCATCCACCTCCACCAGAACATTGTGGACGTGCAGTACTTGTACGGGG
    ACAGGTGGCCGGAGTAGGTGGAGGTGGTCTTGTAACACCTGCACGTCATGAACATGCCCC
                                                ---------------------

    GlySerSerIleAlaSerTrpAlaIleLysTrpGluTyrValValLeuLeuPheLeuLeu
241 TGGGGTCAAGCATCGCGTCCTGGGCCATTAAGTGGGAGTACGTCGTTCTCCTGTTCCTTC
    ACCCCAGTTCGTAGCGCAGGACCCGGTAATTCACCCTCATGCAGCAAGAGGACAAGGAAG
    ------------------------------------------------------------

    LeuAlaAspAlaArgValCysSerCysLeuTrpMetMetLeuLeuIleSerGlnAlaGlu
301 TGCTTGCAGACGCGCGCGTCTGCTCCTGCTTGTGGATGATGCTACTCATATCCCAAGCGG
    ACGAACGTCTGCGCGCGCAGACGAGGACGAACACCTACTACGATGAGTATAGGGTTCGCC
    ---------------Overlap with 14i----------------------------

    AlaAlaLeuGluAsnLeuValIleLeuAsnAlaAlaSerLeuAlaGlyThrHisGlyLeu
361 AGGCGGCTTTGGAGAACCTCGTAATACTTAATGCAGCATCCCTGGCCGGGACGCACGGTC
    TCCGCCGAAACCTCTTGGAGCATTATGAATTACGTCGTAGGGACCGGCCCTGCGTGCCAG
    -------

    Val
421 TTGTATC
    AACATAG
```

FIG. 2-1 Translation of DNA k9-1

```
    GlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGlyTrpGly
  1 CAGGCTGTCCTGAGAGGCTAGCCAGCTGCCGACCCCTTACCGATTTTGACCAGGGCTGGG
    GTCCGACAGGACTCTCCGATCGGTCGACGGCTGGGGAATGGCTAAAACTGGTCCCGACCC

    ProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHisTyrPro
 61 GCCCTATCAGTTATGCCAACGGAAGCGGCCCCGACCAGCGCCCCTACTGCTGGCACTACC
    CGGGATAGTCAATACGGTTGCCTTCGCCGGGGCTGGTCGCGGGGATGACGACCGTGATGG

    ProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCysPheThr
121 CCCCAAAACCTTGCGGTATTGTGCCCGCGAAGAGTGTGTGTGGTCCGGTATATTGCTTCA
    GGGGTTTTGGAACGCCATAACACGGGCGCTTCTCACACACACCAGGCCATATAACGAAGT

    ProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSerTrpGly
181 CTCCCAGCCCCGTGGTGGTGGGAACGACCGACAGGTCGGGCGCGCCCACCTACAGCTGGG
    GAGGGTCGGGGCACCACCACCCTTGCTGGCTGTCCAGCCCGCGCGGGTGGATGTCGACCC

    GluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsnTrpPhe
241 GTGAAAATGATACGGACGTCTTCGTCCTTAACAATACCAGGCCACCGCTGGGCAATTGGT
    CACTTTTACTATGCCTGCAGAAGCAGGAATTGTTATGGTCCGGTGGCGACCCGTTAACCA

    GlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProProCysVal
301 TCGGTTGTACCTGGATGAACTCAACTGGATTCACCAAAGTGTGCGGAGCGCCTCCTTGTG
    AGCCAACATGGACCTACTTGAGTTGACCTAAGTGGTTTCACACGCCTCGCGGAGGAACAC
```

FIG. 2-2

```
    IleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLysHisPro
361 TCATCGGAGGGGCGGCAACAACACCCTGCACTGCCCCACTGATTGCTTCCGCAAGCATC
    AGTAGCCTCCCCGCCCGTTGTTGTGGGACGTGACGGGGTGACTAACGAAGGCGTTCGTAG

    AspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeuValAsp
421 CGGACGCCACATACTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTGCCTGGTCG
    GCCTGCGGTGTATGAGAGCCACGCCGAGGCCAGGGACCTAGTGTGGGTCCACGGACCAGC

                                                   --------------
    TyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLysIleArg
481 ACTACCCGTATAGGCTTTGGCATTATCCTTGTACCATCAACTACACTATATTTAAAATCA
    TGATGGGCATATCCGAAACCGTAATAGGAACATGGTAGTTGATGTGATATAAATTTTAGT

    ------------------------------------------------------------
    MetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArgGlyGlu
541 GGATGTACGTGGGAGGGGTCGAGCACAGGCTGGAAGCTGCCTGCAACTGGACGCGGGGCG
    CCTACATGCACCCTCCCCAGCTCGTGTCCGACCTTCGACGGACGTTGACCTGCGCCCCGC

    ------------------------------------------------------------
    ArgCysAspLeuGluAspArgAspArgSerGluLeuSerProLeuLeuLeuThrThrThr
601 AACGTTGCGATCTGGAAGATAGGGACAGGTCCGAGCTCAGCCCGTTACTGCTGACCACTA
    TTGCAACGCTAGACCTTCTATCCCTGTCCAGGCTCGAGTCGGGCAATGACGACTGGTGAT

    ------------------------------------------------------------
    GlnTrpGlnValLeuProCysSerPheThrThrLeuProAlaLeuSerThrGlyLeuIle
661 CACAGTGGCAGGTCCTCCCGTGTTCCTTCACAACCCTGCCAGCCTTGTCCACCGGCCTCA
    GTGTCACCGTCCAGGAGGGCACAAGGAAGTGTTGGGACGGTCGGAACAGGTGGCCGGAGT
```

FIG. 2-3

```
     -----Overlap with Combined ORF of DNAs 12f through 15e------
        HisLeuHisGlnAsnIleValAspValGlnTyrLeuTyrGlyValGlySerSerIleAla
721 TCCACCTCCACCAGAACATTGTGGACGTGCAGTACTTGTACGGGTGGGGTCAAGCATCG
    AGGTGGAGGTGGTCTTGTAACACCTGCACGTCATGAACATGCCCCACCCCAGTTCGTAGC

    ------------------------------------------------------------
        SerTrpAlaIleLysTrpGluTyrValValLeuLeuPheLeuLeuLeuAlaAspAlaArg
781 CGTCCTGGGCCATTAAGTGGGAGTACGTCGTCCTCCTGTTCCTTCTGCTTGCAGACGCGC
    GCAGGACCCGGTAATTCACCCTCATGCAGCAGGAGGACAAGGAAGACGAACGTCTGCGCG

    ------------------------------------------------------------
        ValCysSerCysLeuTrpMetMetLeuLeuIleSerGlnAlaGluAlaAlaLeuGluAsn
841 GCGTCTGCTCCTGCTTGTGGATGATGCTACTCATATCCCAAGCGGAAGCGGCTTTGGAGA
    CGCAGACGAGGACGAACACCTACTACGATGAGTATAGGGTTCGCCTTCGCCGAAACCTCT

    ------------------------------------------------------------
        LeuValIleLeuAsnAlaAlaSerLeuAlaGlyThrHisGlyLeuValSerPheLeuVal
901 ACCTCGTAATACTTAATGCAGCATCCCTGGCCGGGACGCACGGTCTTGTATCCTTCCTCG
    TGGAGCATTATGAATTACGTCGTAGGGACCGGCCCTGCGTGCCAGAACATAGGAAGGAGC

    ------------------------------------------------------------
        PhePheCysPheAlaTrpTyrLeuLysGlyLysTrpValProGlyAlaValTyrThrPhe
961 TGTTCTTCTGCTTTGCATGGTATCTGAAGGGTAAGTGGGTGCCCGGAGCGGTCTACACCT
    ACAAGAAGACGAAACGTACCATAGACTTCCCATTCACCCACGGGCCTCGCCAGATGTGGA
```

FIG. 2-4

```
     ----------------------------------------------------------
     TyrGlyMetTrpProLeuLeuLeuLeuLeuLeuAlaLeuProGlnArgAlaTyrAlaLeu
1021 TCTACGGGATGTGGCCTCTCCTCCTGCTCCTGTTGGCGTTGCCCCAGCGGGCGTACGCGC
     AGATGCCCTACACCGGAGAGGAGGACGAGGACAACCGCAACGGGGTCGCCCGCATGCGCG

     ----------------------------------------------------------
     AspThrGluValAlaAlaSerCysGlyGlyValValLeuValGlyLeuMetAlaLeuThr
1081 TGGACACGGAGGTGGCCGCGTCGTGTGGCGGTGTTGTTCTCGTCGGGTTGATGGCGCTAA
     ACCTGTGCCTCCACCGGCGCAGCACACCGCCACAACAAGAGCAGCCCAACTACCGCGATT

     ----------------------------------------------------------
     LeuSerProTyrTyrLysArgTyrIleSerTrpCysLeuTrpTrpLeuGlnTyrPheLeu
1141 CTCTGTCACCATATTACAAGCGCTATATCAGCTGGTGCTTGTGGTGGCTTCAGTATTTTC
     GAGACAGTGGTATAATGTTCGCGATATAGTCGACCACGAACACCACCGAAGTCATAAAAG

     ----------------------------------------------------------
     ThrArgValGluAlaGlnLeuHisValTrpIleProProLeuAsnValArgGlyGlyArg
1201 TGACCAGAGTGGAAGCGCAACTGCACGTGTGGATTCCCCCCCTCAACGTCCGAGGGGGGC
     ACTGGTCTCACCTTCGCGTTGACGTGCACACCTAAGGGGGGGAGTTGCAGGCTCCCCCCG

     ----------------------------------------------------------
     AspAlaValIleLeuLeuMetCysAlaValHisProThrLeuValPheAspIleThrLys
1261 GCGACGCTGTCATCTTACTCATGTGTGCTGTACACCCGACTCTGGTATTTGACATCACCA
     CGCTGCGACAGTAGAATGAGTACACACGACATGTGGGCTGAGACCATAAACTGTAGTGGT

     ------------------------------------------
     LeuLeuLeuAlaValPheGlyProLeuTrpIleLeuGlnAla
1321 AATTGCTGCTGGCCGTCTTCGGACCCCTTTGGATTCTTCAAGCCAG
     TTAACGACGACCGGCAGAAGCCTGGGGAAACCTAAGAAGTTCGGTC
```

FIG. 3 Translation of DNA 15e

```
     ----------------------------------------------------------------------
      GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla
  1  CGGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGC
     GCCGCGACCTTTCTCCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCG

     ----------------------Overlap with 26g------------------------------
      AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe
 61  TGCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTT
     ACGCACCCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAA

     ---------------------
      AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla
121  TGCCCCCACACTGTGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGC
     ACGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCG

      ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu
181  CAGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGA
     GTCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCT

      ProLeuAspLeuProProIleIleGlnArgLeu
241  ACCACTTGATCTACCTCCAATCATTCAAAGACTC
     TGGTGAACTAGATGGAGGTTAGTAAGTTTCTGAG
```

**FIG.4** Translation of DNA 13i

```
    ProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSerTrpGly
  1 CTCCCAGCCCCGTGGTGGTGGGAACGACCGACAGGTCGGGCGCGCCTACCTACAGCTGGG
    GAGGGTCGGGGCACCACCACCCTTGCTGGCTGTCCAGCCCGCGCGGATGGATGTCGACCC

    GluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsnTrpPhe
 61 GTGAAAATGATACGGACGTCTTCGTCCTTAACAATACCAGGCCACCGCTGGGCAATTGGT
    CACTTTTACTATGCCTGCAGAAGCAGGAATTGTTATGGTCCGGTGGCGACCCGTTAACCA

    GlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProProCysVal
121 TCGGTTGTACCTGGATGAACTCAACTGGATTCACCAAAGTGTGCGGAGCGCCTCCTTGTG
    AGCCAACATGGACCTACTTGAGTTGACCTAAGTGGTTTCACACGCCTCGCGGAGGAACAC

    IleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLysHisPro
181 TCATCGGAGGGGCGGGCAACAACACCCTGCACTGCCCCACTGATTGCTTCCGCAAGCATC
    AGTAGCCTCCCCGCCCGTTGTTGTGGGACGTGACGGGGTGACTAACGAAGGCGTTCGTAG

    AspAlaThrTyrSerArgCysGlySerGlyProTrpLeuThrProArgCysLeuValAsp
241 CGGACGCCACATACTCTCGGTGCGGCTCCGGTCCCTGGCTCACACCCAGGTGCCTGGTCG
    GCCTGCGGTGTATGAGAGCCACGCCGAGGCCAGGGACCGAGTGTGGGTCCACGGACCAGC
                                                -------------

    TyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLysIleArg
301 ACTACCCGTATAGGCTTTGGCATTATCCTTGTACCATCAACTACACCATATTTAAAATCA
    TGATGGGCATATCCGAAACCGTAATAGGAACATGGTAGTTGATGTGGTATAAATTTTAGT
    ------------------------------------------------------------

    MetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArgGlyGlu
361 GGATGTACGTGGGAGGGGTCGAGCACAGGCTGGAAGCTGCCTGCAACTGGACGCGGGGCG
    CCTACATGCACCCTCCCCAGCTCGTGTCCGACCTTCGACGGACGTTGACCTGCGCCCCGC
    --------------Overlap with 12f------------------------------

    ArgCysAspLeuGluAspArgAspArgSerGluLeuSerProLeuLeuLeuThrThrThr
421 AACGTTGCGATCTGGAAGACAGGACAGGTCCGAGCTCAGCCCGTTACTGCTGACCACTA
    TTGCAACGCTAGACCTTCTGTCCTGTCCAGGCTCGAGTCGGGCAATGACGACTGGTGAT
    ------------------------------------------------------------

    GlnTrpGlnValLeuProCysSerPheThrThrLeuProAlaLeuSerThrGlyLeu
481 CACAGTGGCAGGTCCTCCCGTGTTCCTTCACAACCCTGCCAGCCTTGTCCACCGGCCTCA
    GTGTCACCGTCCAGGAGGGCACAAGGAAGTGTTGGGACGGTCGGAACAGGTGGCCGGAGT
```

Translation of DNA 26j

FIG. 5

```
   LeuPheTyrHisHisLysPheAsnSerSerGlyCysProGluArgLeuAlaSerCysArg
  1 GCTTTTCTATCACCACAAGTTCAACTCTTCAGGCTGTCCTGAGAGGCTAGCCAGCTGCCG
    CGAAAAGATAGTGGTGTTCAAGTTGAGAAGTCCGACAGGACTCTCCGATCGGTCGACGGC

   ProLeuThrAspPheAspGlnGlyTrpGlyProIleSerTyrAlaAsnGlySerGlyPro
 61 ACCCCTTACCGATTTTGACCAGGGCTGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCC
    TGGGGAATGGCTAAAACTGGTCCCGACCCCGGGATAGTCAATACGGTTGCCTTCGCCGGG

   AspGlnArgProTyrCysTrpHisTyrProProLysProCysGlyIleValProAlaLys
121 CGACCAGCGCCCCTACTGCTGGCACTACCCCCCAAAACCTTGCGGTATTGTGCCCGCGAA
    GCTGGTCGCGGGGATGACGACCGTGATGGGGGGTTTTGGAACGCCATAACACGGGCGCTT

                                ---Overlap with 13i---
   SerValCysGlyProValTyrCysPheThrProSerProValValVal
181 GAGTGTGTGTGGTCCGGTATATTGCTTCACTCCCAGCCCCGTGGTGGTGGG
    CTCACACACACCAGGCCATATAACGAAGTGAGGGTCGGGGCACCACCACCC
```

Translation of DNA CA59a FIG. 6

```
    LeuValMetAlaGlnLeuLeuArgIleProGlnAlaIleLeuAspMetIleAlaGlyAla
  1 TTGGTAATGGCTCAGCTGCTCCGGATCCCACAAGCCATCTTGGACATGATCGCTGGTGCT
    AACCATTACCGAGTCGACGAGGCCTAGGGTGTTCGGTAGAACCTGTACTAGCGACCACGA

    HisTrpGlyValLeuAlaGlyIleAlaTyrPheSerMetValGlyAsnTrpAlaLysVal
 61 CACTGGGGAGTCCTGGCGGGCATAGCGTATTTCTCCATGGTGGGGAACTGGGCGAAGGTC
    GTGACCCCTCAGGACCGCCCGTATCGCATAAAGAGGTACCACCCCTTGACCCGCTTCCAG

    LeuValValLeuLeuLeuPheAlaGlyValAspAlaGluThrHisValThrGlyGlySer
121 CTGGTAGTGCTGCTGCTATTTGCCGGCGTCGACGCGGAAACCCACGTCACCGGGGGAAGT
    GACCATCACGACGACGATAAACGGCCGCAGCTGCGCCTTTGGGTGCAGTGGCCCCCTTCA

    AlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLysGlnAsnVal
181 GCCGGCCACACTGTGTCTGGATTTGTTAGCCTCCTCGCACCAGGCGCCAAGCAGAACGTC
    CGGCCGGTGTGACACAGACCTAAACAATCGGAGGAGCGTGGTCCGCGGTTCGTCTTGCAG

    GlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsnCysAsnAsp
241 CAGCTGATCAACACCAACGGCAGTTGGCACCTCAATAGCACGGCCCTGAACTGCAATGAT
    GTCGACTAGTTGTGGTTGCCGTCAACCGTGGAGTTATCGTGCCGGGACTTGACGTTACTA
                                                         -----

    SerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsnSerSerGly
301 AGCCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCACCACAAGTTCAACTCTTCAGGC
    TCGGAGTTGTGGCCGACCAACCGTCCCGAAAAGATAGTGGTGTTCAAGTTGAGAAGTCCG
                                ------Overlap with 26j------------

    ------Overlap with K9-1--------
    CysProGluArgLeuAlaSerCysArgPro
361 TGTCCTGAGAGGCTAGCCAGCTGCCGACCCC
    ACAGGACTCTCCGATCGGTCGACGGCTGGGG
    -------------------------------
```

FIG. 7 Translation of DNA CA84a

```
  GlnGlyCysAsnCysSerIleTyrProGlyHisIleThrGlyHisArgMetAlaTrpAsp
1 CGCAAGGTTGCAATTGCTCTATCTATCCCGGCCATATAACGGGTCACCGCATGGCATGGG
  GCGTTCCAACGTTAACGAGATAGATAGGGCCGGTATATTGCCCAGTGGCGTACCGTACCC

                                                  ---------------------------
  MetMetMetAsnTrpSerProThrThrAlaLeuValMetAlaGlnLeuLeuArgIlePro
61 ATATGATGATGAACTGGTCCCCTACGACGGCGTTGGTAATGGCTCAGCTGCTCCGGATCC
  TATACTACTACTTGACCAGGGGATGCTGCCGCAACCATTACCGAGTCGACGAGGCCTAGG

  ------------------------------------------------------------
  GlnAlaIleLeuAspMetIleAlaGlyAlaHisTrpGlyValLeuAlaGlyIleAlaTyr
121 CACAAGCCATCTTGGACATGATCGCTGGTGCTCACTGGGGAGTCCTGGCGGGCATAGCGT
  GTGTTCGGTAGAACCTGTACTAGCGACCACGAGTGACCCCTCAGGACCGCCCGTATCGCA

  ------------Overlap with CA59a-------------------------------
  PheSerMetValGlyAsnTrpAlaLysValLeuValValLeuLeuLeuPheAlaGlyVal
181 ATTTCTCCATGGTGGGAACTGGGCGAAGGTCCTGGTAGTGCTGCTGCTATTTGCCGGCG
  TAAAGAGGTACCACCCCTTGACCCGCTTCCAGGACCATCACGACGACGATAAACGGCCGC

  ---------------------------
  AspAlaGluThrHisValThrGly
241 TCGACGCGGAAACCCACGTCACCGGGG
  AGCTGCGCCTTTGGGTGCAGTGGCCCC
```

FIG. 8 Translation of DNA CA156e

```
     CysTrpValAlaMetThrProThrValAlaThrArgAspGlyLysLeuProAlaThrGln
   1 GTGTTGGGTGGCGATGACCCCTACGGTGGCCACCAGGGATGGCAAACTCCCCGCGACGCA
     CACAACCCACCGCTACTGGGGATGCCACCGGTGGTCCCTACCGTTTGAGGGGCGCTGCGT

     LeuArgArgHisIleAspLeuLeuValGlySerAlaThrLeuCysSerAlaLeuTyrVal
  61 GCTTCGACGTCACATCGATCTGCTTGTCGGGAGCGCCACCCTCTGTTCGGCCCTCTACGT
     CGAAGCTGCAGTGTAGCTAGACGAACAGCCCTCGCGGTGGGAGACAAGCCGGGAGATGCA

     GlyAspLeuCysGlySerValPheLeuValGlyGlnLeuPheThrPheSerProArgArg
 121 GGGGGACCTATGCGGGTCTGTCTTTCTTGTCGGCCAACTGTTCACCTTCTCTCCCAGGCG
     CCCCCTGGATACGCCCAGACAGAAAGAACAGCCGGTTGACAAGTGGAAGAGAGGGTCCGC

                 ------------------------------------------------
     HisTrpThrThrGlnGlyCysAsnCysSerIleTyrProGlyHisIleThrGlyHisArg
 181 CCACTGGACGACGCAAGGTTGCAATTGCTCTATCTATCCCGGCCATATAACGGGTCACCG
     GGTGACCTGCTGCGTTCCAACGTTAACGAGATAGATAGGGCCGGTATATTGCCCAGTGGC

     --------------------Overlap with CA84a---------------------
     MetAlaTrpAspMetMetMetAsnTrpSerProThrThrAlaLeuValValAlaGlnLeu
 241 CATGGCATGGGATATGATGATGAACTGGTCCCCTACGACGGCGTTGGTAGTGGCTCAGCT
     GTACCGTACCCTATACTACTACTTGACCAGGGGATGCTGCCGCAACCATCACCGAGTCGA

     --------------------
     LeuArgIleProGlnAla
 301 GCTCCGGATCCCACAAGCC
     CGAGGCCTAGGGTGTTCGG
```

FIG. 9 Translation of DNA CA167b

```
   SerThrGlyLeuTyrHisValThrAsnAspCysProAsnSerSerIleValTyrGluAla
  1CTCCACGGGGCTTTACCACGTCACCAATGATTGCCCTAACTCGAGTATTGTGTACGAGGC
   GAGGTGCCCCGAAATGGTGCAGTGGTTACTAACGGGATTGAGCTCATAACACATGCTCCG

   AlaAspAlaIleLeuHisThrProGlyCysValProCysValArgGluGlyAsnAlaSer
 61GGCCGATGCCATCCTGCACACTCCGGGGTGCGTCCCTTGCGTTCGTGAGGGCAACGCCTC
   CCGGCTACGGTAGGACGTGTGAGGCCCCACGCAGGGAACGCAAGCACTCCCGTTGCGGAG

     ----------------------------------------------------------
   ArgCysTrpValAlaMetThrProThrValAlaThrArgAspGlyLysLeuProAlaThr
121GAGGTGTTGGGTGGCGATGACCCCTACGGTGGCCACCAGGGATGGCAAACTCCCCGCGAC
   CTCCACAACCCACCGCTACTGGGGATGCCACCGGTGGTCCCTACCGTTTGAGGGGCGCTG

   ------------------Overlap with CA156e----------------------
   GlnLeuArgArgHisIleAspLeuLeuValGlySerAlaThrLeuCysSerAlaLeuTyr
181GCAGCTTCGACGTCACATCGATCTGCTTGTCGGAGCGCTACCCTCTGTTCGGCCCTCTA
   CGTCGAAGCTGCAGTGTAGCTAGACGAACAGCCCTCGCGATGGGAGACAAGCCGGGAGAT

   ------------------------------
   ValGlyAspLeuCysGlySerValPheLeu
241CGTGGGGGACTTGTGCGGGTCTGTCTTTCTTG
   GCACCCCCTGAACACGCCCAGACAGAAAGAAC
```

FIG.10 Translation of DNA ssCA216a

```
     ArgArgArgSerArgAsnLeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAsp
  1  CCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGATACCCTTACGTGCGGCTTCGCCG
     GGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTATGGGAATGCACGCCGAAGCGGC
     LeuMetGlyTyrIleProLeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAla
 61  ACCTCATGGGGTACATACCGCTCGTCGGCGCCCCTCTTGGAGGCGCTGCCAGGGCCCTGG
     TGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAACCTCCGCGACGGTCCCGGGACC
     HisGlyValArgValLeuGluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCys
121  CGCATGGCGTCCGGGTTCTGGAAGACGGCGTGAACTATGCAACAGGGAACCTTCCTGGTT
     GCGTACCGCAGGCCCAAGACCTTCTGCCGCACTTGATACGTTGTCCCTTGGAAGGACCAA
     SerPheSerIlePheLeuLeuAlaLeuLeuSerCysLeuThrValProAlaSerAlaTyr
181  GCTCTTTCTCTATCTTCCTTCTGGCCCTGCTCTCTTGCTTGACTGTGCCCGCTTCGGCCT
     CGAGAAAGAGATAGAAGGAAGACCGGGACGAGAGAACGAACTGACACGGGCGAAGCCGGA
                         ----------------------------------------
     GlnValArgAsnSerThrGlyLeuTyrHisValThrAsnAspCysProAsnSerSerIle
241  ACCAAGTGCGCAACTCCACGGGGCTTTACCACGTCACCAATGATTGCCCTAACTCGAGTA
     TGGTTCACGCGTTGAGGTGCCCCGAAATGGTGCAGTGGTTACTAACGGGATTGAGCTCAT

     --------------------overlap with CA167b---------------------
     ValTyrGluAlaAlaAspAlaIleLeuHisThrProGlyCysValProCysValArgGlu
301  TTGTGTACGAAGCGGCCGATGCCATCCTGCACACTCCGGGGTGCGTCCCTTGCGTTCGTG
     AACACATGCTTCGCCGGCTACGGTAGGACGTGTGAGGCCCCACGCAGGGAACGCAAGCAC

     ---------------------------------------------
     GlyAsnAlaSerArgCysTrpValAlaMetThrProThrValAla
361  AGGGCAACGCCTCGAGGTGTTGGGTGGCGATGACCCCTACGGTGGCC
     TCCCGTTGCGGAGCTCCACAACCCACCGCTACTGGGGATGCCACCGG
```

FIG.11

```
          Translation of DNA ssCA290a
     LysLysAsnLysArgAsnThrAsnArgArgProGlnAspValLysPheProGlyGlyGly
  1  AAAAAAAAACAAACGTAACACCAACCGTCGCCCACAGGACGTCAAGTTCCCGGGTGGCG
     TTTTTTTTTGTTTGCATTGTGGTTGGCAGCGGGTGTCCTGCAGTTCAAGGGCCCACCGC
     GlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeuGlyValArgAla
 61  GTCAGATCGTTGGTGGAGTTTACTTGTTGCCGCGCAGGGGCCCTAGATTGGGTGTGCGCG
     CAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAACCCACACGCGC
     ThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnProIleProLysAla
121  CGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGCCAGCCTATCCCCAAGG
     GCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCGGTCGGATAGGGGTTCC
     ArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpProLeuTyrGlyAsn
181  CTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCCCTCTATGGCA
     GAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCATGGGAACCGGGGAGATACCGT
     GluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArgProSerTrpGly
241  ATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGGCCTAGCTGGG
     TACTCCCGACGCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCCGGATCGACCC
                ---------------------------------------------------

     ProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAspThrLeuThrCys
301  GCCCCACAGACCCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGATACCCTTACGT
     CGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTATGGGAATGCA
     ------------------------------------------------------------

     GlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeuGlyGlyAlaAla
361  GCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCGGCGCCCCTCTTGGAGGCGCTG
     CGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAACCTCCGCGAC
     ----------------------overlap with CA216a-------------------
     ArgAlaLeuAlaHisGlyValArgValLeuGluAspGlyValAsnTyrAlaThrGlyAsn
421  CCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAAGACGGCGTGAACTATGCAACAGGGA
     GGTCCCGGGACCGCGTACCGCAGGCCCAAGACCTTCTGCCGCACTTGATACGTTGTCCCT
     --------------------------

     LeuProGlyCysSerPheSerThrPhe
481  ACCTTCCTGGTTGCTCTTTCTCTACCTTC
     TGGAAGGACCAACGAGAAAGAGATGGAAG
```

FIG.12-1 Translation of DNA ag30a

```
                                  #MetSerValValGlnProProGlyProProLeu
                     #MetAlaLeuValOP
  1   CGCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCGTGCAGCCTCCAGGACCCCCCC
      GCGTCTTTCGCAGATCGGTACCGCAATCATACTCACAGCACGTCGGAGGTCCTGGGGGGG

        ProGlyGluProAM

 61   TCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAGGACGAC
      AGGGCCCTCTCGGTATCACCAGACGCCTTGGCCACTCATGTGGCCTTAACGGTCCTGCTG

                                  #MetProGlyAspLeuGlyValProProGlnAsp

121   CGGGTCCTTTCTTGGATCAACCCGCTCAATGCCTGGAGATTTGGGCGTGCCCCCGCAAGA
      GCCCAGGAAAGAACCTAGTTGGGCGAGTTACGGACCTCTAAACCCGCACGGGGGCGTTCT

                                                 OP AM GlyAlaCys
       CysAM                                                *

181   CTGCTAGCCGAGTAGTGTTGGGTCGCGAAAGGCCTTGTGGTACTGCCTGATAGGGTGCTT
      GACGATCGGCTCATCACAACCCAGCGCTTTCCGGAACACCATGACGGACTATCCCACGAA
                                          |
        GluCysProGlyArgSerArgArgProCysThrMetSerThrAsnProLysProGlnLys
```

```
241  GCGAGTGCCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCACGAATCCTAAACCTCAAA
     CGCTCACGGGGCCCTCCAGAGCATCTGGCACGTGGTACTCGTGCTTAGGATTTGGAGTTT

       LysAsnLysArgAsnThrAsnArgArgProGlnAspValLysPheProGlyGlyGlyGln

     ------------------------------------------------------------
301  AAAAAAACAAACGTAACACCAACCGTCGCCCACAGGACGTCAAGTTCCCGGGTGGCGGTC
     TTTTTTTGTTTGCATTGTGGTTGGCAGCGGGTGTCCTGCAGTTCAAGGGCCCACCGCCAG

       IleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeuGlyValArgAlaThr

     ------------------------------------------------------------
361  AGATCGTTGGTGGAGTTTACTTGTTGCCGCGCAGGGGCCCTAGATTGGGTGTGCGCGCGA
     TCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAACCCACACGCGCGCT

       ArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnProIleProLysAlaArg

     ------------------------------------------------------------
421  CGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCTATCCCCAAGGCTC
     GCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGATAGGGGTTCCGAG

       ArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpProLeuTyrGlyAsnGlu
```

FIG.12-2

```
     ------------------------overlap with CA290a-----------------
481  GTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCCCTCTATGGCAATG
     CAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCATGGGAACCGGGGAGATACCGTTAC

       GlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArgProSerTrpGlyPro

     ------------------------------------------------------------
541  AGGGCTGCGGGTGGGCGGATGGCTCCTGTCTCCCCGTGGCTCTCGGCCTAGCTGGGGCC
     TCCCGACGCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCCGGATCGACCCCGG

       ThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAspThrLeuThrCysGly

601  CCACAGACCCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGATACCCTTACGTGCG
     GGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTATGGGAATGCACGC

       Phe

     -----
661  GCTTC
     CGAAG
```

* = Start of long HCV ORF
| = Putative first amino acid of large HCV polyprotein
# = Putative small encoded peptides (that may play a translational regulatory role)

FIG.12-3

## FIG. 13

Translation of DNA CA205a

```
    ValLeuGlyArgGluArgProCysGlyThrAlaOP AM GlyAlaCysGluCysProGly
  1 GTCTTGGGTCGCGAAAGGCCTTGTGGTACTGCCTGATAGGGTGCTTGCGAGTGCCCCGGG
    CAGAACCCAGCGCTTTCCGGAACACCATGACGGACTATCCCACGAACGCTCACGGGGCCC

                              *                -----------------
    ArgSerArgArgProCysThrMetSerThrAsnProLysProGlnArgLysThrLysArg
 61 AGGTCTCGTAGACCGTGCACCATGAGCACGAATCCTAAACCTCAAAGAAAAACCAAACGT
    TCCAGAGCATCTGGCACGTGGTACTCGTGCTTAGGATTTGGAGTTTCTTTTTGGTTTGCA

    ------------------------------------------------------------
    AsnThrAsnArgArgProGlnAspValLysPheProGlyGlyGlyGlnIleValGlyGly
121 AACACCAACCGTCGCCCACAGGACGTCAAGTTCCCGGGTGGCGGTCAGATCGTTGGTGGA
    TTGTGGTTGGCAGCGGGTGTCCTGCAGTTCAAGGGCCCACCGCCAGTCTAGCAACCACCT

    ------------------------------------------------------------
    ValTyrLeuLeuProArgArgGlyProArgLeuGlyValArgAlaThrArgLysThrSer
181 GTTTACTTGTTGCCGCGCAGGGGCCCTAGATTGGGTGTGCGCGCGACGAGAAAGACTTCC
    CAAATGAACAACGGCGCGTCCCCGGGATCTAACCCACACGCGCGCTGCTCTTTCTGAAGG

    ------------------overlap with CA290a-----------------------
    GluArgSerGlnProArgGlyArgArgGlnProIleProLysAlaArgArgProGluGly
241 GAGCGGTCGCAACCTCGAGGTAGACGTCAGCCTATCCCCAAGGCTCGTCGGCCCGAGGGC
    CTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGATAGGGGTTCCGAGCAGCCGGGCTCCCG

    ---------------------------------------------------
    ArgThrTrpAlaGlnProGlyTyrProTrpProLeuTyrGlyAsnGluGlyCys
301 AGGACCTGGGCTCAGCCCGGGTACCCTTGGCCCCTCTATGGCAATGAGGGCTGCG
    TCCTGGACCCGAGTCGGGCCCATGGGAACCGGGGAGATACCGTTACTCCCGACGC
```

* = putative initiator methionine codon

FIG.14 Translation of DNA 18g

```
                                                    ----------------------
      #ProProOP
      #SerThrMetAsnHisSerProValArgAsnTyrCysLeuHisAlaGluSerValAM Pro
     #LeuHisHisGluSerLeuProCysGluGluLeuLeuSerSerArgArgLysArgLeuAla
  1   CTCCACCATGAATCACTCCCCTGTGAGGAACTACTGTCTTCACGCAGAAAGCGTCTAGCC
      GAGGTGGTACTTAGTGAGGGGACACTCCTTGATGACAGAAGTGCGTCTTTCGCAGATCGG
      ------------------------------------------------------------
                #MetSerValValGlnProProGlyProProLeuProGlyGluProAM
      MetAlaLeuValOP
 61   ATGGCGTTAGTATGAGTGTCGTGCAGCCTCCAGGACCCCCCCTCCCGGGAGAGCCATAGT
      TACCGCAATCATACTCACAGCACGTCGGAGGTCCTGGGGGGGAGGGCCCTCTCGGTATCA

      ------------------------------------------------------------
121   GGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAGGACGACCGGGTCCTTTCTTGGATC
      CCAGACGCCTTGGCCACTCATGTGGCCTTAACGGTCCTGCTGGCCCAGGAAAGAACCTAG
      ---------------overlap with ag30a---------------------------

               #MetProGlyAspLeuGlyValProProGlnAspCysAM
181   AACCCGCTCAATGCCTGGAGATTTGGGCGTGCCCCCGCAAGACTGCTAGCCGAGTAGTGT
      TTGGGCGAGTTACGGACCTCTAAACCCGCACGGGGGCGTTCTGACGATCGGCTCATCACA
      ------------------------------------------------------------
                                    OP AM GlyAlaCysGluCysProGlyArgSer
                                             *
241   TGGGTCGCGAAAGGCCTTGTGGTACTGCCTGATAGGGTGCTTGCGAGTGCCCCGGGAGGT
      ACCCAGCGCTTTCCGGAACACCATGACGGACTATCCCACGAACGCTCACGGGGCCCTCCA

      --------
        ArgArg
301   CTCGTAGA
      GAGCATCT
```

* = Start of long HCV ORF
\# = Putative small encoded peptides (that may play a translational regulatory role)

FIG. 15 Translation of DNA 16jh

```
     ------Overlap with 15e ------------------------------
     GlyAlaCysTyrSerIleGluProLeuAspLeuProProIleIleGlnArgLeuHisGly
   1 GGGGCCTGCTACTCCATAGAACCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGC
     CCCCGGACGATGAGGTATCTTGGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCG

     LeuSerAlaPheSerLeuHisSerTyrSerProGlyGluIleAsnArgValAlaAlaCys
  61 CTCAGCGCATTTTCACTCCACAGTTACTCTCCAGGTGAAATTAATAGGGTGGCCGCATGC
     GAGTCGCGTAAAAGTGAGGTGTCAATGAGAGGTCCACTTTAATTATCCCACCGGCGTACG

                                         Gly*
                                         G
     LeuArgLysLeuGlyValProProLeuArgAlaTrpArgHisArgAlaArgSerValArg
 121 CTCAGAAAACTTGGGGTACCGCCCTTGCGAGCTTGGAGACACCGGCCCGGAGCGTCCGC
     GAGTCTTTTGAACCCCATGGCGGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCG

     AlaArgLeuLeuAlaArgGlyGlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrp
 181 GCTAGGCTTCTGGCCAGAGGAGGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGG
     CGATCCGAAGACCGGTCTCCTCCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACC

     AlaValArgThrLysLeuLys
 241 GCAGTAAGAACAAAGCTCAAAC
     CGTCATTCTTGTTTCGAGTTTG
```

* = nucleotide heterogeneity

FIG. 16 Translation of DNA 6k

```
     -----------------------Overlap with 16jh-------------------
     GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
  1  GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA
     CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTCGAGTTT
     -
     LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
 61  CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC
     GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG

     SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
121  AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCCGCTGGATCTGGTTTTGC
     TCGCCCCCTCTGTAAATAGTGTCGCACAGAGTACGGGCCGGGGCGACCTAGACCAAAACG

181  CC
     GG
```

Translation of DNA p131jh

```
      --------------Overlap with 6k------------------
       TyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCysLeuLeuLeuLeuAla
  1   TTATCACAGCGTGTCTCATGCCCGGCCCCGCTGGATCTGGTTTTGCCTACTCCTGCTTGC
      AATAGTGTCGCACAGAGTACGGGCCGGGGCGACCTAGACCAAAACGGATGAGGACGAACG

       AlaGlyValGlyIleTyrLeuLeuProAsnArgOP
 61   TGCAGGGGTAGGCATCTACCTCCTCCCCAACCGATGAAGGTTGGGGTAAACACTCCGGCC
      ACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTACTTCCAACCCCATTTGTGAGGCCGG

121   T
      A
```

FIG.17

FIG. 18-1

```
-341   GCCAGCCCCCTGATGGGGGCGA
       CGGTCGGGGGACTACCCCCGCT

-319   CACTCCACCATGAATCACTCCCCTGTGAGGAACTACTGTCTTCACGCAGAAAGCGTCTAG
       GTGAGGTGGTACTTAGTGAGGGGACACTCCTTGATGACAGAAGTGCGTCTTTCGCAGATC

-259   CCATGGCGTTAGTATGAGTGTCGTGCAGCCTCCAGGACCCCCCCTCCCGGGAGAGCCATA
       GGTACCGCAATCATACTCACAGCACGTCGGAGGTCCTGGGGGGGAGGGCCCTCTCGGTAT

-199   GTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAGGACGACCGGGTCCTTTCTTGGA
       CACCAGACGCCTTGGCCACTCATGTGGCCTTAACGGTCCTGCTGGCCCAGGAAAGAACCT

-139   TCAACCCGCTCAATGCCTGGAGATTTGGGCGTGCCCCCGCAAGACTGCTAGCCGAGTAGT
       AGTTGGGCGAGTTACGGACCTCTAAACCCGCACGGGGGCGTTCTGACGATCGGCTCATCA

- 79   GTTGGGTCGCGAAAGGCCTTGTGGTACTGCCTGATAGGGTGCTTGCGAGTGCCCCGGGAG
       CAACCCAGCGCTTTCCGGAACACCATGACGGACTATCCCACGAACGCTCACGGGGCCCTC

- 19   GTCTCGTAGACCGTGCACC
       CAGAGCATCTGGCACGTGG

                                Arg   Thr
       MetSerThrAsnProLysProGlnLysLysAsnLysArgAsnThrAsnArgArgProGln
  1    ATGAGCACGAATCCTAAACCTCAAAAAAAAACAAACGTAACACCAACCGTCGCCCACAG
       TACTCGTGCTTAGGATTTGGAGTTTTTTTTTGTTTGCATTGTGGTTGGCAGCGGGTGTC

       AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArg
 61    GACGTCAAGTTCCCGGGTGGCGGTCAGATCGTTGGTGGAGTTTACTTGTTGCCGCGCAGG
       CTGCAGTTCAAGGGCCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCC
```

FIG. 18-2

```
     GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly
121  GGCCCTAGATTGGGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGT
     CCGGGATCTAACCCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCA

     ArgArgGlnProIleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGly
181  AGACGTCAGCCTATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGG
     TCTGCAGTCGGATAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCC

     TyrProTrpProLeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerPro
241  TACCCTTGGCCCCTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCC
     ATGGGAACCGGGGAGATACCGTTACTCCCGACGCCCACCCGCCCTACCGAGGACAGAGGG

     ArgGlySerArgProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGly
301  CGTGGCTCTCGGCCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTCGCGCAATTTGGGT
     GCACCGAGAGCCGGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCA

     LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuVal
361  AAGGTCATCGATACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTC
     TTCCAGTAGCTATGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAG

     GlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp
421  GGCGCCCCTCTTGGAGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAAGAC
     CCGCGGGGAGAACCTCCGCGACGGTCCCGGGACCGCGTACCGCAGGCCCAAGACCTTCTG
```

FIG. 18-3

```
                                                  Thr
     GlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla
481  GGCGTGAACTATGCAACAGGGAACCTTCCTGGTTGCTCTTTCTCTATCTTCCTTCTGGCC
     CCGCACTTGATACGTTGTCCCTTGGAAGGACCAACGAGAAAGAGATAGAAGGAAGACCGG

     LeuLeuSerCysLeuThrValProAlaSerAlaTyrGlnValArgAsnSerThrGlyLeu
541  CTGCTCTCTTGCTTGACTGTGCCCGCTTCGGCCTACCAAGTGCGCAACTCCACGGGGCTT
     GACGAGAGAACGAACTGACACGGGCGAAGCCGGATGGTTCACGCGTTGAGGTGCCCCGAA

     TyrHisValThrAsnAspCysProAsnSerSerIleValTyrGluAlaAlaAspAlaIle
601  TACCACGTCACCAATGATTGCCCTAACTCGAGTATTGTGTACGAGGCGGCCGATGCCATC
     ATGGTGCAGTGGTTACTAACGGGATTGAGCTCATAACACATGCTCCGCCGGCTACGGTAG

     LeuHisThrProGlyCysValProCysValArgGluGlyAsnAlaSerArgCysTrpVal
661  CTGCACACTCCGGGGTGCGTCCCTTGCGTTCGTGAGGGCAACGCCTCGAGGTGTTGGGTG
     GACGTGTGAGGCCCCACGCAGGGAACGCAAGCACTCCCGTTGCGGAGCTCCACAACCCAC

     AlaMetThrProThrValAlaThrArgAspGlyLysLeuProAlaThrGlnLeuArgArg
721  GCGATGACCCCTACGGTGGCCACCAGGGATGGCAAACTCCCCGCGACGCAGCTTCGACGT
     CGCTACTGGGGATGCCACCGGTGGTCCCTACCGTTTGAGGGGCGCTGCGTCGAAGCTGCA

     HisIleAspLeuLeuValGlySerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu
781  CACATCGATCTGCTTGTCGGGAGCGCCACCCTCTGTTCGGCCCTCTACGTGGGGGACCTA
     GTGTAGCTAGACGAACAGCCCTCGCGGTGGGAGACAAGCCGGGAGATGCACCCCCTGGAT

     CysGlySerValPheLeuValGlyGlnLeuPheThrPheSerProArgArgHisTrpThr
841  TGCGGGTCTGTCTTTCTTGTCGGCCAACTGTTCACCTTCTCTCCCAGGCGCCACTGGACG
     ACGCCCAGACAGAAAGAACAGCCGGTTGACAAGTGGAAGAGAGGGTCCGCGGTGACCTGC
```

FIG. 18-4

```
     ThrGlnGlyCysAsnCysSerIleTyrProGlyHisIleThrGlyHisArgMetAlaTrp
901  ACGCAAGGTTGCAATTGCTCTATCTATCCCGGCCATATAACGGGTCACCGCATGGCATGG
     TGCGTTCCAACGTTAACGAGATAGATAGGGCCGGTATATTGCCCAGTGGCGTACCGTACC

                                           Val
     AspMetMetMetAsnTrpSerProThrThrAlaLeuValMetAlaGlnLeuLeuArgIle
961  GATATGATGATGAACTGGTCCCCTACGACGGCGTTGGTAATGGCTCAGCTGCTCCGGATC
     CTATACTACTACTTGACCAGGGGATGCTGCCGCAACCATTACCGAGTCGACGAGGCCTAG

     ProGlnAlaIleLeuAspMetIleAlaGlyAlaHisTrpGlyValLeuAlaGlyIleAla
1021 CCACAAGCCATCTTGGACATGATCGCTGGTGCTCACTGGGGAGTCCTGGCGGGCATAGCG
     GGTGTTCGGTAGAACCTGTACTAGCGACCACGAGTGACCCCTCAGGACCGCCCGTATCGC

     TyrPheSerMetValGlyAsnTrpAlaLysValLeuValValLeuLeuLeuPheAlaGly
1081 TATTTCTCCATGGTGGGGAACTGGGCGAAGGTCCTGGTAGTGCTGCTGCTATTTGCCGGC
     ATAAAGAGGTACCACCCCTTGACCCGCTTCCAGGACCATCACGACGACGATAAACGGCCG

     ValAspAlaGluThrHisValThrGlyGlySerAlaGlyHisThrValSerGlyPheVal
1141 GTCGACGCGGAAACCCACGTCACCGGGGGAAGTGCCGGCCACACTGTGTCTGGATTTGTT
     CAGCTGCGCCTTTGGGTGCAGTGGCCCCCTTCACGGCCGGTGTGACACAGACCTAAACAA

     SerLeuLeuAlaProGlyAlaLysGlnAsnValGlnLeuIleAsnThrAsnGlySerTrp
1201 AGCCTCCTCGCACCAGGCGCCAAGCAGAACGTCCAGCTGATCAACACCAACGGCAGTTGG
     TCGGAGGAGCGTGGTCCGCGGTTCGTCTTGCAGGTCGACTAGTTGTGGTTGCCGTCAACC
```

FIG. 18-5

```
     HisLeuAsnSerThrAlaLeuAsnCysAsnAspSerLeuAsnThrGlyTrpLeuAlaGly
1261 CACCTCAATAGCACGGCCCTGAACTGCAATGATAGCCTCAACACCGGCTGGTTGGCAGGG
     GTGGAGTTATCGTGCCGGGACTTGACGTTACTATCGGAGTTGTGGCCGACCAACCGTCCC

     LeuPheTyrHisHisLysPheAsnSerSerGlyCysProGluArgLeuAlaSerCysArg
1321 CTTTTCTATCACCACAAGTTCAACTCTTCAGGCTGTCCTGAGAGGCTAGCCAGCTGCCGA
     GAAAAGATAGTGGTGTTCAAGTTGAGAAGTCCGACAGGACTCTCCGATCGGTCGACGGCT

     ProLeuThrAspPheAspGlnGlyTrpGlyProIleSerTyrAlaAsnGlySerGlyPro
1381 CCCCTTACCGATTTTGACCAGGGCTGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCCC
     GGGGAATGGCTAAAACTGGTCCCGACCCCGGGATAGTCAATACGGTTGCCTTCGCCGGGG

     AspGlnArgProTyrCysTrpHisTyrProProLysProCysGlyIleValProAlaLys
1441 GACCAGCGCCCCTACTGCTGGCACTACCCCCCAAAACCTTGCGGTATTGTGCCCGCGAAG
     CTGGTCGCGGGGATGACGACCGTGATGGGGGGTTTTGGAACGCCATAACACGGGCGCTTC

     SerValCysGlyProValTyrCysPheThrProSerProValValValGlyThrThrAsp
1501 AGTGTGTGTGGTCCGGTATATTGCTTCACTCCCAGCCCCGTGGTGGTGGGAACGACCGAC
     TCACACACACCAGGCCATATAACGAAGTGAGGGTCGGGGCACCACCACCCTTGCTGGCTG

     ArgSerGlyAlaProThrTyrSerTrpGlyGluAsnAspThrAspValPheValLeuAsn
1561 AGGTCGGGCGCGCCCACCTACAGCTGGGGTGAAAATGATACGGACGTCTTCGTCCTTAAC
     TCCAGCCCGCGCGGGTGGATGTCGACCCCACTTTTACTATGCCTGCAGAAGCAGGAATTG

     AsnThrArgProProLeuGlyAsnTrpPheGlyCysThrTrpMetAsnSerThrGlyPhe
```

FIG. 18-6

```
1621  AATACCAGGCCACCGCTGGGCAATTGGTTCGGTTGTACCTGGATGAACTCAACTGGATTC
      TTATGGTCCGGTGGCGACCCGTTAACCAAGCCAACATGGACCTACTTGAGTTGACCTAAG

      ThrLysValCysGlyAlaProProCysValIleGlyGlyAlaGlyAsnAsnThrLeuHis
1681  ACCAAAGTGTGCGGAGCGCCTCCTTGTGTCATCGGAGGGGCGGGCAACAACACCCTGCAC
      TGGTTTCACACGCCTCGCGGAGGAACACAGTAGCCTCCCCGCCCGTTGTTGTGGGACGTG

      CysProThrAspCysPheArgLysHisProAspAlaThrTyrSerArgCysGlySerGly
1741  TGCCCCACTGATTGCTTCCGCAAGCATCCGGACGCCACATACTCTCGGTGCGGCTCCGGT
      ACGGGGTGACTAACGAAGGCGTTCGTAGGCCTGCGGTGTATGAGAGCCACGCCGAGGCCA

         Ile
      ProTrpLeuThrProArgCysLeuValAspTyrProTyrArgLeuTrpHisTyrProCys
1801  CCCTGGATCACACCCAGGTGCCTGGTCGACTACCCGTATAGGCTTTGGCATTATCCTTGT
      GGGACCTAGTGTGGGTCCACGGACCAGCTGATGGGCATATCCGAAACCGTAATAGGAACA

      ThrIleAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyValGluHisArgLeu
1861  ACCATCAACTACACCATATTTAAAATCAGGATGTACGTGGAGGGGTCGAACACAGGCTG
      TGGTAGTTGATGTGGTATAAATTTTAGTCCTACATGCACCCTCCCCAGCTTGTGTCCGAC

      GluAlaAlaCysAsnTrpThrArgGlyGluArgCysAspLeuGluAspArgAspArgSer
1921  GAAGCTGCCTGCAACTGGACGCGGGGCGAACGTTGCGATCTGGAAGACAGGGACAGGTCC
      CTTCGACGGACGTTGACCTGCGCCCCGCTTGCAACGCTAGACCTTCTGTCCCTGTCCAGG

      GluLeuSerProLeuLeuLeuThrThrThrGlnTrpGlnValLeuProCysSerPheThr
1981  GAGCTCAGCCCGTTACTGCTGACCACTACACAGTGGCAGGTCCTCCCGTGTTCCTTCACA
      CTCGAGTCGGGCAATGACGACTGGTGATGTGTCACCGTCCAGGAGGGCACAAGGAAGTGT
```

FIG. 18-7

```
      ThrLeuProAlaLeuSerThrGlyLeuIleHisLeuHisGlnAsnIleValAspValGln
2041  ACCCTACCAGCCTTGTCCACCGGCCTCATCCACCTCCACCAGAACATTGTGGACGTGCAG
      TGGGATGGTCGGAACAGGTGGCCGGAGTAGGTGGAGGTGGTCTTGTAACACCTGCACGTC

      TyrLeuTyrGlyValGlySerSerIleAlaSerTrpAlaIleLysTrpGluTyrValVal
2101  TACTTGTACGGGGTGGGGTCAAGCATCGCGTCCTGGGCCATTAAGTGGGAGTACGTCGTT
      ATGAACATGCCCCACCCCAGTTCGTAGCGCAGGACCCGGTAATTCACCCTCATGCAGCAA

      LeuLeuPheLeuLeuLeuAlaAspAlaArgValCysSerCysLeuTrpMetMetLeuLeu
2161  CTCCTGTTCCTTCTGCTTGCAGACGCGCGCGTCTGCTCCTGCTTGTGGATGATGCTACTC
      GAGGACAAGGAAGACGAACGTCTGCGCGCGCAGACGAGGACGAACACCTACTACGATGAG

      IleSerGlnAlaGluAlaAlaLeuGluAsnLeuValIleLeuAsnAlaAlaSerLeuAla
2221  ATATCCCAAGCGGAGGCGGCTTTGGAGAACCTCGTAATACTTAATGCAGCATCCCTGGCC
      TATAGGGTTCGCCTCCGCCGAAACCTCTTGGAGCATTATGAATTACGTCGTAGGGACCGG

      GlyThrHisGlyLeuValSerPheLeuValPhePheCysPheAlaTrpTyrLeuLysGly
2281  GGGACGCACGGTCTTGTATCCTTCCTCGTGTTCTTCTGCTTTGCATGGTATTTGAAGGGT
      CCCTGCGTGCCAGAACATAGGAAGGAGCACAAGAAGACGAAACGTACCATAAACTTCCCA

      LysTrpValProGlyAlaValTyrThrPheTyrGlyMetTrpProLeuLeuLeuLeuLeu
2341  AAGTGGGTGCCCGGAGCGGTCTACACCTTCTACGGGATGTGGCCTCTCCTCCTGCTCCTG
      TTCACCCACGGGCCTCGCCAGATGTGGAAGATGCCCTACACCGGAGAGGAGGACGAGGAC
```

FIG. 18-8

```
     LeuAlaLeuProGlnArgAlaTyrAlaLeuAspThrGluValAlaAlaSerCysGlyGly
2401 TTGGCGTTGCCCCAGCGGGCGTACGCGCTGGACACGGAGGTGGCCGCGTCGTGTGGCGGT
     AACCGCAACGGGGTCGCCCGCATGCGCGACCTGTGCCTCCACCGGCGCAGCACACCGCCA

     ValValLeuValGlyLeuMetAlaLeuThrLeuSerProTyrTyrLysArgTyrIleSer
2461 GTTGTTCTCGTCGGGTTGATGGCGCTGACTCTGTCACCATATTACAAGCGCTATATCAGC
     CAACAAGAGCAGCCCAACTACCGCGACTGAGACAGTGGTATAATGTTCGCGATATAGTCG

                             (Asn)
     TrpCysLeuTrpTrpLeuGlnTyrPheLeuThrArgValGluAlaGlnLeuHisValTrp
2521 TGGTGCTTGTGGTGGCTTCAGTATTTTCTGACCAGAGTGGAAGCGCAACTGCACGTGTGG
     ACCACGAACACCACCGAAGTCATAAAAGACTGGTCTCACCTTCGCGTTGACGTGCACACC

     IleProProLeuAsnValArgGlyGlyArgAspAlaValIleLeuLeuMetCysAlaVal

2581 ATTCCCCCCTCAACGTCCGAGGGGGGCGCGACGCCGTCATCTTACTCATGTGTGCTGTA
     TAAGGGGGGAGTTGCAGGCTCCCCCCGCGCTGCGGCAGTAGAATGAGTACACACGACAT

     HisProThrLeuValPheAspIleThrLysLeuLeuLeuAlaValPheGlyProLeuTrp
2641 CACCCGACTCTGGTATTTGACATCACCAAATTGCTGCTGGCCGTCTTCGGACCCCTTTGG
     GTGGGCTGAGACCATAAACTGTAGTGGTTTAACGACGACCGGCAGAAGCCTGGGGAAACC

     IleLeuGlnAlaSerLeuLeuLysValProTyrPheValArgValGlnGlyLeuLeuArg
2701 ATTCTTCAAGCCAGTTTGCTTAAAGTACCCTACTTTGTGCGCGTCCAAGGCCTTCTCCGG
     TAAGAAGTTCGGTCAAACGAATTTCATGGATGAAACACGCGCAGGTTCCGGAAGAGGCC
```

FIG. 18-9

```
     PheCysAlaLeuAlaArgLysMetIleGlyGlyHisTyrValGlnMetValIleIleLys
2761 TTCTGCGCGTTAGCGCGGAAGATGATCGGAGGCCATTACGTGCAAATGGTCATCATTAAG
     AAGACGCGCAATCGCGCCTTCTACTAGCCTCCGGTAATGCACGTTTACCAGTAGTAATTC

     LeuGlyAlaLeuThrGlyThrTyrValTyrAsnHisLeuThrProLeuArgAspTrpAla
2821 TTAGGGGCGCTTACTGGCACCTATGTTTATAACCATCTCACTCCTCTTCGGGACTGGGCG
     AATCCCCGCGAATGACCGTGGATACAAATATTGGTAGAGTGAGGAGAAGCCCTGACCCGC

     HisAsnGlyLeuArgAspLeuAlaValAlaValGluProValValPheSerGlnMetGlu
2881 CACAACGGCTTGCGAGATCTGGCCGTGGCTGTAGAGCCAGTCGTCTTCTCCCAAATGGAG
     GTGTTGCCGAACGCTCTAGACCGGCACCGACATCTCGGTCAGCAGAAGAGGGTTTACCTC

     ThrLysLeuIleThrTrpGlyAlaAspThrAlaAlaCysGlyAspIleIleAsnGlyLeu
2941 ACCAAGCTCATCACGTGGGGGGCAGATACCGCCGCGTGCGGTGACATCATCAACGGCTTG
     TGGTTCGAGTAGTGCACCCCCCGTCTATGGCGGCGCACGCCACTGTAGTAGTTGCCGAAC

     ProValSerAlaArgArgGlyArgGluIleLeuLeuGlyProAlaAspGlyMetValSer
3001 CCTGTTTCCGCCCGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCC
     GGACAAAGGCGGGCGTCCCCGGCCCTCTATGACGAGCCCGGTCGGCTACCTTACCAGAGG

     LysGlyTrpArgLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeu
3061 AAGGGGTGGAGGTTGCTGGCGCCCATCACGGCGTACGCCCAGCAGACAAGGGGCCTCCTA
     TTCCCCACCTCCAACGACCGCGGGTAGTGCCGCATGCGGGTCGTCTGTTCCCCGGAGGAT

     GlyCysIleIleThrSerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGln
3121 GGGTGCATAATCACCAGCCTAACTGGCCGGGACAAAAACCAAGTGGAGGGTGAGGTCCAG
     CCCACGTATTAGTGGTCGGATTGACCGGCCCTGTTTTTGGTTCACCTCCCACTCCAGGTC
```

FIG. 18-10

```
      IleValSerThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThr
3181  ATTGTGTCAACTGCTGCCCAAACCTTCCTGGCAACGTGCATCAATGGGGTGTGCTGGACT
      TAACACAGTTGACGACGGGTTTGGAAGGACCGTTGCACGTAGTTACCCCACACGACCTGA

      ValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMet
3241  GTCTACCACGGGGCCGGAACGAGGACCATCGCGTCACCCAAGGGTCCTGTCATCCAGATG
      CAGATGGTGCCCCGGCCTTGCTCCTGGTAGCGCAGTGGGTTCCCAGGACAGTAGGTCTAC

                                             Ser      Thr
      TyrThrAsnValAspGlnAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeu
3301  TATACCAATGTAGACCAAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCCGCTCATTG
      ATATGGTTACATCTGGTTCTGGAACACCCGACCGGGCGAGGCGTTCCATCGGCGAGTAAC

      ThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIle
3361  ACACCCTGCACTTGCGGCTCCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGTCATT
      TGTGGGACGTGAACGCCGAGGAGCCTGGAAATGGACCAGTGCTCCGTGCGGCTACAGTAA

      ProValArgArgArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyr
3421  CCCGTGCGCCGGCGGGGTGATAGCAGGGGCAGCCTGCTGTCGCCCCGGCCCATTTCCTAC
      GGGCACGCGGCCGCCCCACTATCGTCCCCGTCGGACGACAGCGGGGCCGGGTAAAGGATG

      LeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePhe
3481  TTGAAAGGCTCCTCGGGGGGTCCGCTGTTGTGCCCCGCGGGGCACGCCGTGGGCATATTT
      AACTTTCCGAGGAGCCCCCCAGGCGACAACACGGGGCGCCCCGTGCGGCACCCGTATAAA

      ArgAlaAlaValCysThrArgGlyValAlaLysAlaValAspPheIleProValGluAsn
3541  AGGGCCGCGGTGTGCACCCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGAAC
```

FIG. 18-11

```
      TCCCGGCGCCACACGTGGGCACCTCACCGATTCCGCCACCTGAAATAGGGACACCTCTTG

      LeuGluThrThrMetArgSerProValPheThrAspAsnSerSerProProValValPro
3601  CTAGAGACAACCATGAGGTCCCCGGTGTTCACGGATAACTCCTCTCCACCAGTAGTGCCC
      GATCTCTGTTGGTACTCCAGGGGCCACAAGTGCCTATTGAGGAGAGGTGGTCATCACGGG

      GlnSerPheGlnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysVal
3661  CAGAGCTTCCAGGTGGCTCACCTCCATGCTCCCACAGGCAGCGGCAAAAGCACCAAGGTC
      GTCTCGAAGGTCCACCGAGTGGAGGTACGAGGGTGTCCGTCGCCGTTTTCGTGGTTCCAG

      ProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAla
3721  CCGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCA
      GGCCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTGGGGAGACAACGACGT

                                                Leu
      ThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThr
3781  ACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGATCCTAACATCAGGACC
      TGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTAGGATTGTAGTCCTGG

      GlyValArgThrIleThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeu
3841  GGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACCTACGGCAAGTTCCTT
      CCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGGATGCCGTTCAAGGAA

      AlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSer
3901  GCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCACTCC
      CGGCTGCCGCCCACGAGCCCCCCGCGAATACTGTATTATTAAACACTGCTCACGGTGAGG
```

FIG. 18-12

```
                         (Val)
      ThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGly
3961  ACGGATGCCACATCCATCTTGGGCATCGGCACTGTCCTTGACCAAGCAGAGACTGCGGGG
      TGCCTACGGTGTAGGTAGAACCCGTAGCCGTGACAGGAACTGGTTCGTCTCTGACGCCCC

      AlaArgLeuValValLeuAlaThrAlaThrProProGlySerValThrValProHisPro
4021  GCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCCATCCC
      CGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAGTGACACGGGGTAGGG

      AsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIle
4081  AACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTTTACGGCAAGGCTATC
      TTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAAATGCCGTTCCGATAG

      ProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCys
4141  CCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTCTGTCATTCAAAGAAGAAGTGC
      GGGGAGCTTCATTAGTTCCCCCCCTCTGTAGAGTAGAAGACAGTAAGTTTCTTCTTCACG

      AspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaValAlaTyrTyrArgGly
4201  GACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTGGCCTACTACCGCGGT
      CTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCACCGGATGATGGCGCCA

      LeuAspValSerValIleProThrSerGlyAspValValValValAlaThrAspAlaLeu
4261  CTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTCGTCGTGGCAACCGATGCCCTC
      GAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCACCGTTGGCTACGGGAG
```

FIG. 18-13

```
                                            Tyr
      MetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnThrCysValThrGln
4321  ATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCACCCAG
      TACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTATGCACACAGTGGGTC

                                   (Ser)
      ThrValAspPheSerLeuAspProThrPheThrIleGluThrIleThrLeuProGlnAsp
4381  ACAGTCGATTTCAGCCTTGACCCTACCTTCACCATTGAGACAATCACGCTCCCCCAGGAT
      TGTCAGCTAAAGTCGGAACTGGGATGGAAGTGGTAACTCTGTTAGTGCGAGGGGGTCCTA

      AlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArg
4441  GCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAGCCAGGCATCTACAGA
      CGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTCGGTCCGTAGATGTCT

      PheValAlaProGlyGluArgProSerGlyMetPheAspSerSerValLeuCysGluCys
4501  TTTGTGGCACCGGGGGAGCGCCCCTCCGGCATGTTCGACTCGTCCGTCCTCTGTGAGTGC
      AAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGGCAGGAGACACTCACG

      TyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArg
4561  TATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACTACAGTTAGGCTACGA
      ATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGATGTCAATCCGATGCT

      AlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGly
4621  GCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGGAGGGC
      CGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTGGTAGAACTTAAAACCCTCCCG

      ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGly
4681  GTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAGACAAAGCAGAGTGGG
      CAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTCTGTTTCGTCTCACCC
```

FIG. 18-14

```
      GluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaPro
4741  GAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCCT
      CTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGATCCCGAGTTCGGGGA

      ProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGly
4801  CCCCCATCGTGGGACCAGATGTGGAAGTGTTTGATTCGCCTCAAGCCCACCCTCCATGGG
      GGGGGTAGCACCCTGGTCTACACCTTCACAAACTAAGCGGAGTTCGGGTGGGAGGTACCC

      ProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIleThrLeuThrHisPro
4861  CCAACACCCCTGCTATACAGACTGGGCGCTGTTCAGAATGAAATCACCCTGACGCACCCA
      GGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAGTGGGACTGCGTGGGT

      ValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluValValThrSerThrTrp
4921  GTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTCGTCACGAGCACCTGG
      CAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAGCAGTGCTCGTGGACC

      ValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysVal
4981  GTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTGTCAACAGGCTGCGTG
      CACGAGCAACCGCCGCAGGACCGACGAAACCGGCGCATAACGGACAGTTGTCCGACGCAC

      ValIleValGlyArgValValLeuSerGlyLysProAlaIleIleProAspArgGluVal
5041  GTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATACCTGACAGGGAAGTC
      CAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTATGGACTGTCCCTTCAG

      LeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeuProTyrIleGluGln
5101  CTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTACCGTACATCGAGCAA
      GAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAATGGCATGTAGCTCGTT
```

FIG. 18-15

```
      GlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeuLeuGlnThrAlaSer
5161  GGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTCCTGCAGACCGCGTCC
      CCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAGGACGTCTGGCGCAGG

      ArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGlnLysLeuGluThrPhe
5221  CGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAAAAACTCGAGACCTTC
      GCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTTTTTGAGCTCTGGAAG

      TrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeuAlaGlyLeuSerThr
5281  TGGGCGAAGCATATGTGGAACTTCATCAGTGGGATACAATACTTGGCGGGCTTGTCAACG
      ACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAACCGCCCGAACAGTTGC

      LeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAlaAlaValThrSerPro
5341  CTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCTGCTGTCACCAGCCCA
      GACGGACCATTGGGGCGGTAACGAAGTAACTACCGAAAATGTCGACGACAGTGGTCGGGT

      LeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrpValAlaAlaGlnLeu
5401  CTAACCACTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGGGTGGCTGCCCAGCTC
      GATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCCACCCACCGACGGGTCGAG

      AlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAlaGlyAlaAlaIleGly
5461  GCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCTGGCGCCGCCATCGGC

      CGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGACCGCGGCGGTAGCCG
```

FIG. 18-16

```
      SerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyrGlyAlaGlyValAla
5521  AGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTATGGCGCGGGCGTGGCG
      TCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATACCGCGCCCGCACCGC

                                                    (Gly)
      GlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSerThrGluAspLeuVal
5581  GGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCCACGGAGGACCTGGTC
      CCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGGTGCCTCCTGGACCAG

      AsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGlyValValCysAlaAla
5641  AATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGCGTGGTCTGTGCAGCA
      TTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCGCACCAGACACGTCGT

      IleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMetAsnArgLeuIle
5701  ATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGGATGAACCGGCTGATA
      TATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACCTACTTGGCCGACTAT

      AlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrValProGluSerAspAla
5761  GCCTTCGCCTCCCGGGGGAACCATGTTTCCCCCACGCACTACGTGCCGGAGAGCGATGCA
      CGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCACGGCCTCTCGCTACGT

                              (HisCys)
      AlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGlnLeuLeuArgArgLeu
5821  GCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAGCTCCTGAGGCGACTG
      CGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTCGAGGACTCCGCTGAC
```

FIG. 18-17

```
     HisGlnTrpIleSerSerGluCysThrThrProCysSerGlySerTrpLeuArgAspIle
5881 CACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCCTGGCTAAGGGACATC
     GTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGGACCGATTCCCTGTAG

     TrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeuLysAlaLysLeuMet
5941 TGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTAAAAGCTAAGCTCATG
     ACCCTGACCTATACGCTCCACAACTCGCTGAAATTCTGGACCGATTTTCGATTCGAGTAC

     ProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyrLysGlyValTrpArg
6001 CCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTATAAGGGGGTCTGGCGA
     GGTGTCGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCATATTCCCCCAGACCGCT

    (Val)
     GlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIleThrGlyHisValLys
6061 GTGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATCACTGGACATGTCAAA
     CACCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAGTGACCTGTACAGTTT

     AsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMetTrpSerGlyThrPhe
6121 AACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATGTGGAGTGGGACCTTC
     TTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGTACACCTCACCCTGGAAG

     ProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAlaProAsnTyrThrPhe
6181 CCCATTAATGCCTACACCACGGGCCCCTGTACCCCCCTTCCTGCGCCGAACTACACGTTC
     GGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGCGGCTTGATGTGCAAG
```

FIG. 18-18

```
     AlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGlnValGlyAspPheHis
6241 GCGCTATGGAGGGTGTCTGCAGAGGAATATGTGGAGATAAGGCAGGTGGGGGACTTCCAC
     CGCGATACCTCCCACAGACGTCTCCTTATACACCTCTATTCCGTCCACCCCCTGAAGGTG

     TyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGlnValProSerProGlu
6301 TACGTGACGGGTATGACTACTGACAATCTCAAATGCCCGTGCCAGGTCCCATCGCCCGAA
     ATGCACTGCCCATACTGATGACTGTTAGAGTTTACGGGCACGGTCCAGGGTAGCGGGCTT

     PhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaProProCysLysProLeu
6361 TTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCCCCCTGCAAGCCCTTG
     AAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGGGGGACGTTCGGGAAC

     LeuArgGluGluValSerPheArgValGlyLeuHisGluTyrProValGlySerGlnLeu

6421 CTGCGGGAGGAGGTATCATTCAGAGTAGGACTCCACGAATACCCGGTAGGGTCGCAATTA
     GACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGCCATCCCAGCGTTAAT

     ProCysGluProGluProAspValAlaValLeuThrSerMetLeuThrAspProSerHis
6481 CCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTCACTGATCCCTCCCAT
     GGAACGCTCGGGCTTGGCCTGCACCGGCACAACTGCAGGTACGAGTGACTAGGGAGGGTA

     IleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerProProSerValAlaSer
6541 ATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGATCACCCCCCTCTGTGGCCAGC
     TATTGTCGTCTCCGCCGGCCCGCTTCCAACCGCTCCCCTAGTGGGGGGAGACACCGGTCG
```

FIG. 18-19

```
      SerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCysThrAlaAsnHisAsp
6601  TCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGCACCGCTAACCATGAC
      AGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGTTGAACGTGGCGATTGGTACTG

      SerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGlnGluMetGlyGlyAsn
6661  TCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAGGAGATGGGCGGCAAC
      AGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTCCTCTACCCGCCGTTG

      IleThrArgValGluSerGluAsnLysValValIleLeuAspSerPheAspProLeuVal
6721  ATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCCTTCGATCCGCTTGTG
      TAGTGGTCCCAACTCAGTCTTTTGTTTCACCACTAAGACCTGAGGAAGCTAGGCGAACAC

      AlaGluGluAspGluArgGluIleSerValProAlaGluIleLeuArgLysSerArgArg
6781  GCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTGCGGAAGTCTCGGAGA
      CGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGACGCCTTCAGAGCCTCT

      PheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnProProLeuValGluThr
6841  TTCGCCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTATAACCCCCCGCTAGTGGAGACG
      AAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGGGGCGATCACCTCTGC

      TrpLysLysProAspTyrGluProProValValHisGlyCysProLeuProProProLys
6901  TGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGTCCGCTTCCACCTCCAAAG
      ACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACAGGCGAAGGTGGAGGTTTC

      SerProProValProProProArgLysLysArgThrValValLeuThrGluSerThrLeu
6961  TCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTCACTGAATCAACCCTA
      AGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAGTGACTTAGTTGGGAT
```

FIG. 18-20

```
                      (Ser)
      SerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSerSerThrSerGlyIle
7021  TCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCCTCAACTTCCGGCATT
      AGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGGAGTTGAAGGCCGTAA

      ThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGlyCysProProAspSer
7081  ACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCCTTCTGGCTGCCCCCCCGACTCC
      TGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCGACGGGGGGGCTGAGG

             (PheAla)
      AspAlaGluSerTyrSerSerMetProProLeuGluGlyGluProGlyAspProAspLeu
7141  GACGCTGAGTCCTATTCCTCCATGCCCCCCCTGGAGGGGGAGCCTGGGGATCCGGATCTT
      CTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGACCCCTAGGCCTAGAA

      SerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGluAspValValCysCys
7201  AGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAGGATGTCGTGTGCTGC
      TCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTCCTACAGCACACGACG

      SerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAlaAlaGluGluGlnLys
7261  TCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCCGCGGAAGAACAGAAA
      AGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGGCGCCTTCTTGTCTTT

      LeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsnLeuValTyrSerThr
7321  CTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAATTTGGTGTATTCCACC
      GACGGGTAGTTACGTGATTCGTTGAGCAACGATGCAGTGGTGTTAAACCACATAAGGTGG

      ThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAspArgLeuGlnValLeu
```

FIG. 18-21

```
7381  ACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGACAGACTGCAAGTTCTG
      TGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAACTGTCTGACGTTCAAGAC

      AspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAlaSerLysValLysAla
7441  GACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCGTCAAAAGTGAAGGCT
      CTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTCGTCGCCGCAGTTTTCACTTCCGA

        (Phe)
      AsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHisSerAlaLysSerLys
7501  AACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCCACACTCAGCCAAATCCAAG
      TTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTGAGTCGGTTTAGGTTC

      PheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAlaValThrHisIleAsn
7561  TTTGGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGAAAGGCCGTAACCCACATCAAC
      AAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGGCATTGGGTGTAGTTG

      SerValTrpLysAspLeuLeuGluAspAsnValThrProIleAspThrThrIleMetAla
7621  TCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGACACTACCATCATGGCT
      AGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTGTGATGGTAGTACCGA

      LysAsnGluValPheCysValGlnProGluLysGlyGlyArgLysProAlaArgLeuIle
7681  AAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAGCCAGCTCGTCTCATC
      TTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCCAGCATTCGGTCGAGCAGAGTAG

      ValPheProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspValValThr
7741  GTGTTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTGTACGACGTGGTTACA
      CACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAACATGCTGCACCAATGT
```

FIG. 18-22

```
     LysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyrSerProGlyGlnArg
7801 AAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATACTCACCAGGACAGCGG
     TTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATGAGTGGTCCTGTCGCC

     ValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMetGlyPheSerTyrAsp
7861 GTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATGGGGTTCTCGTATGAT
     CAACTTAAGGAGCACGTTCGCACCTTCAGGTTCTTTTGGGGTTACCCCAAGAGCATACTA

     ThrArgCysPheAspSerThrValThrGluSerAspIleArgThrGluGluAlaIleTyr
7921 ACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACGGAGGAGGCAATCTAC
     TGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGCCTCCTCCGTTAGATG

     GlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSerLeuThrGluArgLeu
7981 CAATGTTGTGACCTCGACCCCCAAGCCCGCGTGGCCATCAAGTCCCTCACCGAGAGGCTT
     GTTACAACACTGGAGCTGGGGGTTCGGGCGCACCGGTAGTTCAGGGAGTGGCTCTCCGAA

                                (Gly)
     TyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGlyTyrArgArgCysArg
8041 TATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGCTATCGCAGGTGCCGC
     ATACAACCCCCGGGAGAATGGTTAAGTTCCCCCCTCTTGACGCCGATAGCGTCCACGGCG

     AlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCysTyrIleLysAlaArg
8101 GCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGCTACATCAAGGCCCGG
     CGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACGATGTAGTTCCGGGCC
```

FIG. 18-23

```
      AlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuValCysGlyAspAspLeu
8161  GCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTGTGTGGCGACGACTTA
      CGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCACACACCGCTGCTGAAT

      ValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSerLeuArgAlaPheThr
8221  GTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGCCTGAGAGCCTTCACG
      CAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCGGACTCTCGGAAGTGC

      GluAlaMetThrArgTyrSerAlaProProGlyAspProProGlnProGluTyrAspLeu
8281  GAGGCTATGACCAGGTACTCCGCCCCCCCTGGGGACCCCCCACAACCAGAATACGACTTG
      CTCCGATACTGGTCCATGAGGCGGGGGGGACCCCTGGGGGGTGTTGGTCTTATGCTGAAC

      GluLeuIleThrSerCysSerSerAsnValSerValAlaHisAspGlyAlaGlyLysArg
8341  GAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCACGACGGCGCTGGAAAGAGG
      CTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTGCCGCGACCTTTCTCC

      ValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAlaAlaTrpGluThrAla
8401  GTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCTGCGTGGGAGACAGCA
      CAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGACGCACCCTCTGTCGT

      ArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPheAlaProThrLeuTrp
8461  AGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTTGCCCCCACACTGTGG
      TCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAACGGGGGTGTGACACC
```

FIG. 18-24

```
      AlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAlaArgAspGlnLeuGlu
8521  GCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCCAGGACCAGCTTGAA
      CGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGGTCCCTGGTCGAACTT

      GlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGluProLeuAspLeuPro
8581  CAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAACCACTTGATCTACCT
      GTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTTGGTGAACTAGATGGA

      ProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHisSerTyrSerProGly
8641  CCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCACAGTTACTCTCCAGGT
      GGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTGTCAATGAGAGGTCCA

      GluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValProProLeuArgAlaTrp
8701  GAAATTAATAGGGTGGCCGCATGCCTCAGAAAACTTGGGGTACCGCCCTTGCGAGCTTGG
      CTTTAATTATCCCACCGGCGTACGGAGTCTTTTGAACCCCATGGCGGGAACGCTCGAACC

      Gly
      ArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGlyGlyArgAlaAlaIle
8761  AGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGAGGCAGGGCTGCCATA
      TCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCTCCGTCCCGACGGTAT

      CysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLysLeuThrProIleAla
8821  TGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAACTCACTCCAATAGCG
      ACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTCGAGTTTGAGTGAGGTTATCGC
```

FIG. 18-25

```
     AlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyrSerGlyGlyAspIle
8881 GCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTACAGCGGGGGAGACATT
     CGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATGTCGCCCCCTCTGTAA

                                                      (Pro)
     TyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCysLeuLeuLeuLeuAla
8941 TATCACAGCGTGTCTCATGCCCGGCCCCGCTGGATCTGGTTTTGCCTACTCCTGCTTGCT
     ATAGTGTCGCACAGAGTACGGGCCGGGGCGACCTAGACCAAAACGGATGAGGACGAACGA

     AlaGlyValGlyIleTyrLeuLeuProAsnArgOP
9001 GCAGGGGTAGGCATCTACCTCCTCCCCAACCGATGAAGGTTGGGGTAAACACTCCGGCCT
     CGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTACTTCCAACCCCATTTGTGAGGCCGGA
```

:42.16.XT1
GGTAGGGTCAAGGCTGAAATCGACTGTCTGCTTCTTTGGAGAAAGTGGTG

:42.17.XT1
ATCCTGGGGGAGCGTGATTGTCTCAATGGTCTTCTTTGGAGAAAGTGGTG

:42.18.XT1
AGTCCTGCCCCGACGTTGAGTGCGGGAGACCTTCTTTGGAGAAAGTGGTG

:42.19.XT1
CACAAATCTGTAGATGCCTGGCTTCCCCCTCTTCTTTGGAGAAAGTGGTG

:42.20.XT1
GTCGAACATGCCGGAGGGGCGCTCCCCCGGCTTCTTTGGAGAAAGTGGTG

:42.21.LLA2C
GCCTGCGTCATAGCACTCACAGAGGACGGATTAGGCATAGGACCCGTGTC

:42.22.LLA2C
AGTCTCGGCGGGCGTGAGCTCATACCAAGCTTAGGCATAGGACCCGTGTC

:42.23.LLA2C
CGGGGTGTTCATGTACGCTCGTAGCCTAACTTAGGCATAGGACCCGTGTC

:42.24.LLA2C
AAATTCAAGATGGTCCTGGCACACGGGAAGTTAGGCATAGGACCCGTGTC

:42.25.LLA2C
TATATGAGTGAGGCCTGTAAAGACGCCCTCTTAGGCATAGGACCCGTGTC

:42.26.LLA2C
ACTCTGCTTTGTCTGGGATAGAAAGTGGGCTTAGGCATAGGACCCGTGTC

:42.27.LLA2C
TTGGTACGCTACCAGGTAAGGAAGGTTCTCTTAGGCATAGGACCCGTGTC

:42.28.LLA2C
GGGAGGGGCTTGAGCCCTAGCGCACACGGTTTAGGCATAGGACCCGTGTC

:42.29.LLA2C
AATCAAACACTTCCACATCTGGTCCCACGATTAGGCATAGGACCCGTGTC

:42.30.LLA2C
GGGTGTTGGCCCATGGAGGGTGGGCTTGAGTTAGGCATAGGACCCGTGTC

:42.31.LLA2C
TTCATTCTGAACAGCGCCCAGTCTGTATAGTTAGGCATAGGACCCGTGTC

FIG. 19-1

:42.XT1.1
TCCTCACAGGGGAGTGATTCATGGTGGAGTCTTCTTTGGAGAAAGTGGTG

:42.XT1.2
ATGGCTAGACGCTTTCTGCGTGAAGACAGTCTTCTTTGGAGAAAGTGGTG

:42.XT1.3
TCCTGGAGGCTGCACGACACTCATACTAACCTTCTTTGGAGAAAGTGGTG

:42.XT1.4
CGCAGACCACTATGGCTCTCCCGGGAGGGGCTTCTTTGGAGAAAGTGGTG

:42.XT1.5
TCGTCCTGGCAATTCCGGTGTACTCACCGGCTTCTTTGGAGAAAGTGGTG

:42.LLA2C.6
GCATTGAGCGGGTTGATCCAAGAAAGGACCTTAGGCATAGGACCCGTGTC

:42.LLA2C.7
AGCAGTCTTGCGGGGGCACGCCCAAATCTCTTAGGCATAGGACCCGTGTC

:42.LLA2C.8
ACAAGGCCTTTCGCGACCCAACACTACTCGTTAGGCATAGGACCCGTGTC

:42.LLA2C.9
GGGGCACTCGCAAGCACCCTATCAGGCAGTTTAGGCATAGGACCCGTGTC

:42.LLA2.10
CGTGCTCATGGTGCACGGTCTACGAGACCTTTAGGCATAGGACCCGTGTC

:42.LLA2C.11
GTTACGTTTGTTTTTTTTTGAGGTTTAGGTTAGGCATAGGACCCGTGTC

:42.LLA2C.12
CGGGAACTTGACGTCCTGTGGGCGACGGTTTTAGGCATAGGACCCGTGTC

:42.LLA2C.13
CAAGTAAACTCCACCAACGATCTGACCGCCTTAGGCATAGGACCCGTGTC

:42.LLA2C.14
GCGCACACCCAATCTAGGGCCCCTGCGCGGTTAGGCATAGGACCCGTGTC

:42.LLA2C.15
AGGTTGCGACCGCTCGGAAGTCTTTCTCGTTTAGGCATAGGACCCGTGTC

FIG. 19-2

:42.32.XT1
ATGTTGGGATGGGGCACAGTGACGGAGCCCCTTCTTTGGAGAAAGTGGTG

:42.33.XT1
ATCTCTCCGGTGGTGGACAGAGCAACCTCCCTTCTTTGGAGAAAGTGGTG

:42.34.XT1
ACTTCGAGGGGGATAGCCTTGCCGTAAAAACTTCTTTGGAGAAAGTGGTG

:42.35.XT1
TGACAGAAGATGAGATGTCTCCCCCCCTTGCTTCTTTGGAGAAAGTGGTG

:42.36.LLA2C
TTTGCGGCGAGTTCGTCGCACTTCTTCTTTTTAGGCATAGGACCCGTGTC

:42.37.LLA2C
TAGGCCACGGCATTGATGCCCAATGCGACCTTAGGCATAGGACCCGTGTC

:42.38.LLA2C
GTCGGGATGACGGACACGTCAAGACCGCGGTTAGGCATAGGACCCGTGTC

:42.39.LLA2C
GCATCGGTTGCCACGACGACAACATCGCCGTTAGGCATAGGACCCGTGTC

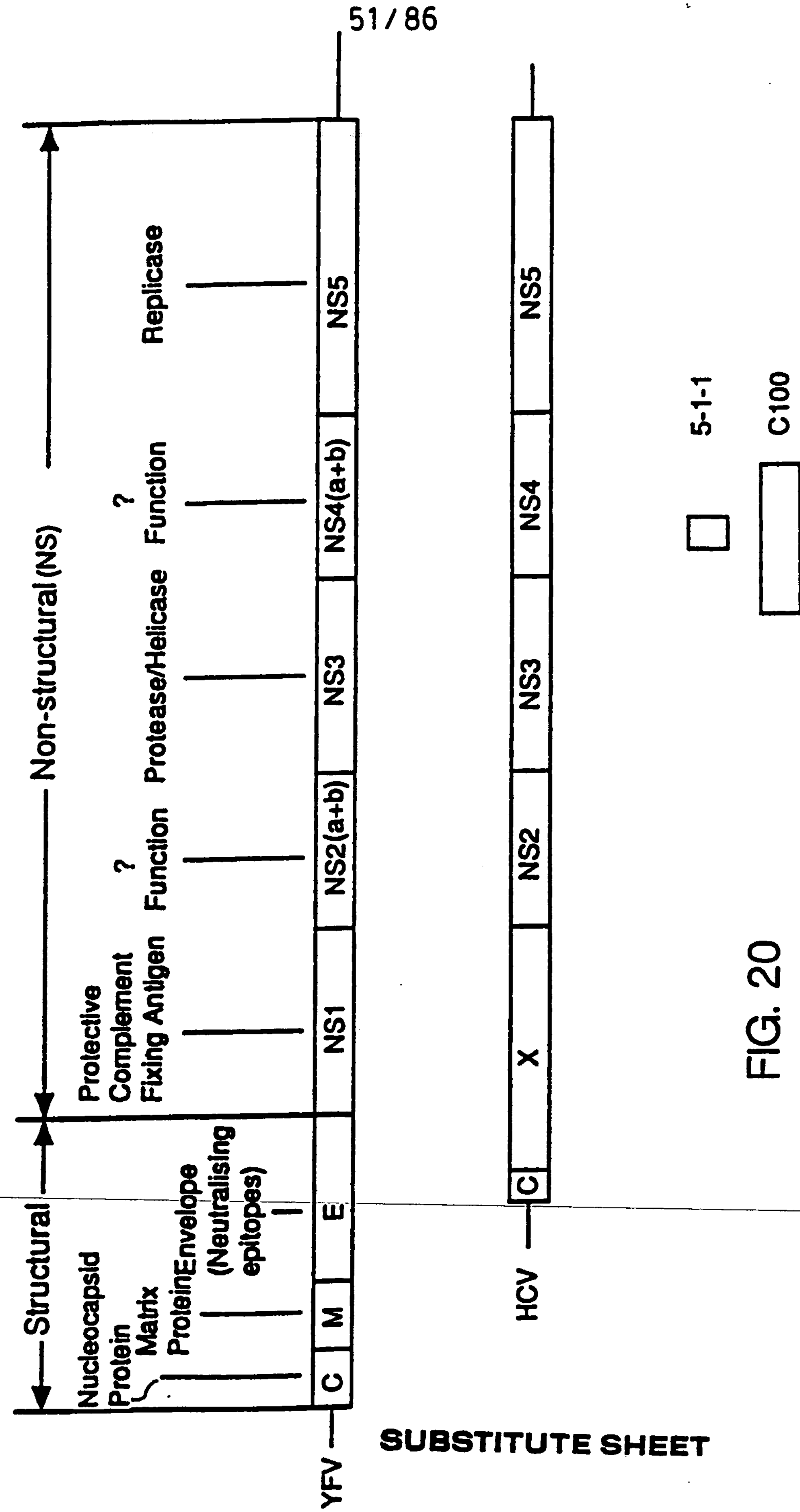
:42.40.LLA2C
GAGTCGAAGTCGCCGGTATAGCCGGTCATGTTAGGCATAGGACCCGTGTC

:42.41.LLA2C
GTCTGGGTGACACACGTATTGCAGTCTATCTTAGGCATAGGACCCGTGTC

:42.42.LLA2C
ATGGTGAAGGTAGGGTCAAGGCTGAAATCGTTAGGCATAGGACCCGTGTC

:42.43.LLA2C
GAGACAGCATCCTGGGGGAGCGTGATTGTCTTAGGCATAGGACCCGTGTC

FIG. 19-3



Structural
Non-structural (NS)
Nucleocapsid Protein
Matrix Protein
Envelope (Neutralising epitopes)
Protective Complement Fixing Antigen
? Function
Protease/Helicase
? Function
Replicase
YFV
C
M
E
NS1
NS2(a+b)
NS3
NS4(a+b)
NS5
HCV
C
X
NS2
NS3
NS4
NS5
5-1-1
C100


FIG. 20

FIG. 22 Translation of DNA 81

```
    SerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPheAspGluMet
  1 GTCCGGGAAGCCGGCAATCATACCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGAT
    CAGGCCCTTCGGCCGTTAGTATGGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTA

    GluGluCysSerGlnHisLeuProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPhe
 61 GGAAGAGTGCTCTCAGCACTTACCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTT
    CCTTCTCACGAGAGTCGTGAATGGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAA

    LysGlnLysAlaLeuGlyLeuLeuGlnThrAlaSerArgGlnAlaGluValIleAlaPro
121 CAAGCAGAAGGCCCTCGGCCTCCTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCC
    GTTCGTCTTCCGGGAGCCGGAGGACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGG

    AlaValGlnThrAsnTrpGlnLysLeuGluThrPheTrpAlaLysHisMetTrpAsnPhe
181 TGCTGTCCAGACCAACTGGCAAAAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACTT
    ACGACAGGTCTGGTTGACCGTTTTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAA

    IleSerGlyIleGlnTyrLeuAlaGlyLeuSerThrLeuProGlyAsnProAlaIleAla
241 CATCAGTGGGATACAATACTTGGCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGC
    GTAGTCACCCTATGTTATGAACCGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACG

    SerLeuMetAlaPheThrAlaAlaValThrSerProLeuThrThrSerGln
301 TTCATTGATGGCTTTTACAGCTGCTGTCACCAGCCCACTAACCACTAGCCAAA
    AAGTAACTACCGAAAATGTCGACGACAGTGGTCGGTGATTGGTGATCGGTTT
```

FIG. 23 Translation of DNA 36

```
   AspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAla
1  GATGCCCACTTTCTATCCCAGACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCG
   CTACGGGTGAAAGATAGGGTCTGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGC

   TyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAspGlnMetTrp
61 TACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCCTCCCCCATCGTGGACCAGATGTGG
   ATGGTTCGGTGGCACACGCGATCCCGAGTTCGGGGAGGGGGTAGCACCCTGGTCTACACC

    LysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeuTyrArgLeu
121 AAGTGTTTGATTCGCCTCAAGCCCACCCTCCATGGCCAACACCCCTGCTATACAGACTG
    TTCACAAACTAAGCGGAGTTCGGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGAC

    GlyAlaValGlnAsnGluIleThrLeuThrHisProValThrLysTyrIleMetThrCys
181 GGCGCTGTTCAGAATGAAATCACCCTGACGCACCCAGTCACCAAATACATCATGACATGC
    CCGCGACAAGTCTTACTTTAGTGGGACTGCGTGGGTCAGTGGTTTATGTAGTACTGTACG

    MetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGlyValLeuAla
241 ATGTCGGCCGACCTGGAGGTCGTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCT
    TACAGCCGGCTGGACCTCCAGCAGTGCTCGTGGACCCACGAGCAACCGCCGCAGGACCGA

    AlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArgValValLeu
301 GCTTTGGCCGCGTATTGCCTGTCAACAGGCTGCGTGGTCATAGTGGGCAGGTCGTCTTG
    CGAAACCGGCGCATAACGGACAGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAAC

    -----------Overlap with 81--------------------
    SerGlyLysProAlaIleIleProAspArgGluValLeuTyrArg
361 TCCGGGAAGCCGGCAATCATACCTGACAGGGAAGTCCTCTACCGAG
    AGGCCCTTCGGCCGTTAGTATGGACTGTCCCTTCAGGAGATGGCTC
```

FIG. 24 Translation of DNA 37b

```
    LeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaValAlaTyrTyrArgGlyLeuAsp
  1 CTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTGGCCTACTACCGCGGTCTTGAC
    GAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCACCGGATGATGGCGCCAGAACTG

    ValSerValIleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThr
 61 GTGTCCGTCATCCCGACCAGCGGCGATGTTGTCGTCGTGGCAACCGATGCCCTCATGACC
    CACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCACCGTTGGCTACGGGAGTACTGG

    GlyTyrThrGlyAspPheAspSerValIleAspTyrAsnThrCysValThrGlnThrVal
121 GGCTATACCGGCGACTTCGACTCGGTGATAGACTACAATACGTGTGTCACCCAGACAGTC
    CCGATATGGCCGCTGAAGCTGAGCCACTATCTGATGTTATGCACACAGTGGGTCTGTCAG

                                -----------------------Overlap with
    AspPheSerLeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaVal
181 GATTTCAGCCTTGACCCTACCTTCACCATTGAGACAATCACGCTCCCCCAGGATGCTGTC
    CTAAAGTCGGAACTGGGATGGAAGTGGTAACTCTGTTAGTGCGAGGGGGTCCTACGACAG

    clone 35-------------------
    SerArgThrGlnArgArgGlyArgThr
241 TCCCGCACTCAACGTCGGGGCAGGACTG
    AGGGCGTGAGTTGCAGCCCCGTCCTGAC
```

FIG. 25A

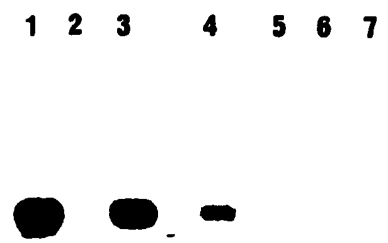



FIG. 25B

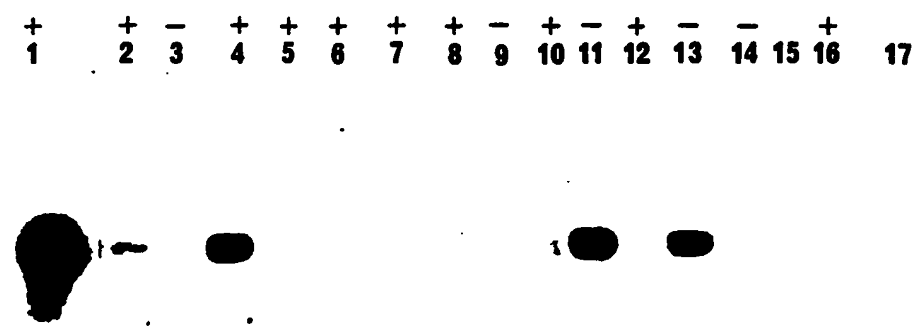


FIG. 26A

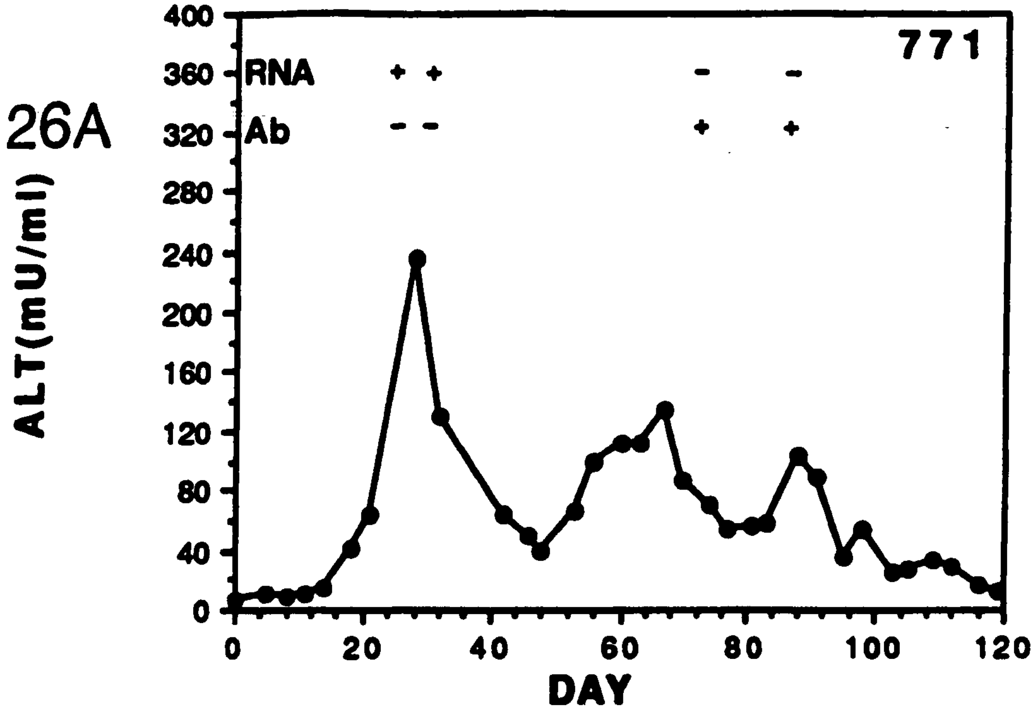
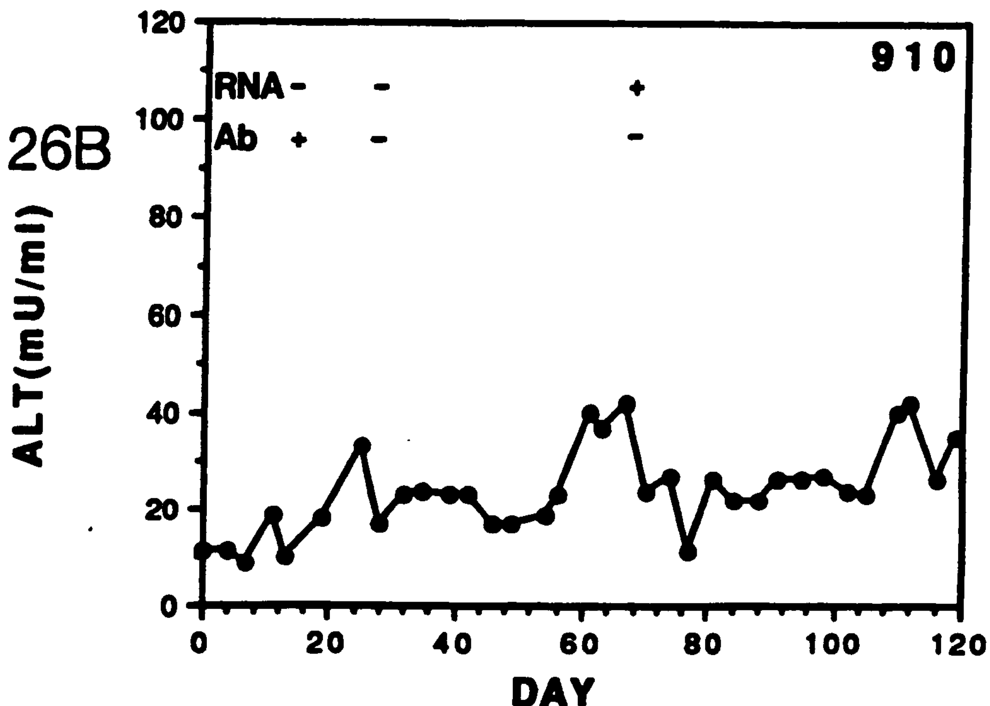



FIG. 26B




1 2 3 4 a b c

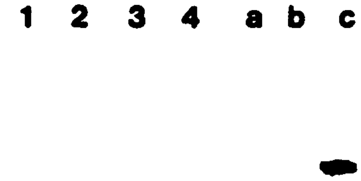


FIG. 26B'

Translation of DNA CA84a

```
      GlnGlyCysAsnCysSerIleTyrProGlyHisIleThrGlyHisArgMetAlaTrpAsp
  1   CGCAAGGTTGCAATTGCTCTATCTATCCCGGCCATATAACGGGTCACCGCATGGCATGGG
      GCGTTCCAACGTTAACGAGATAGATAGGGCCGGTATATTGCCCAGTGGCGTACCGTACCC

                                              ------------------------------
      MetMetMetAsnTrpSerProThrThrAlaLeuValMetAlaGlnLeuLeuArgIlePro
 61   ATATGATGATGAACTGGTCCCCTACGACGGCGTTGGTAATGGCTCAGCTGCTCCGGATCC
      TATACTACTACTTGACCAGGGGATGCTGCCGCAACCATTACCGAGTCGACGAGGCCTAGG

      ------------------------------------------------------------
      GlnAlaIleLeuAspMetIleAlaGlyAlaHisTrpGlyValLeuAlaGlyIleAlaTyr
121   CACAAGCCATCTTGGACATGATCGCTGGTGCTCACTGGGGAGTCCTGGCGGGCATAGCGT
      GTGTTCGGTAGAACCTGTACTAGCGACCACGAGTGACCCCTCAGGACCGCCCGTATCGCA

      ------------Overlap with CA59a-------------------------------
      PheSerMetValGlyAsnTrpAlaLysValLeuValValLeuLeuLeuPheAlaGlyVal
181   ATTTCTCCATGGTGGGAACTGGCGAAGGTCCTGGTAGTGCTGCTGCTATTTGCCGGCG
      TAAAGAGGTACCACCCTTGACCCGCTTCCAGGACCATCACGACGACGATAAACGGCCGC

      ------------------------
      AspAlaGluThrHisValThrGly
241   TCGACGCGGAAACCCACGTCACCGGGG
      AGCTGCGCCTTTGGGTGCAGTGGCCCC
```

FIG. 27

```
     AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle
1    GGCTTACATGTCCAAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAAT
     CCGAATGTACAGGTTCCGAGTACCCTAGCTAGGATTGTAGTCCTGGCCCCACTCTTGTTA

     ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys
61   TACCACTGGCAGCCCCATCACGTACTCCACCTACGGCAAGTTCCTTGCCGACGGCGGGTG
     ATGGTGACCGTCGGGGTAGTGCATGAGGTGGATGCCGTTCAAGGAACGGCTGCCGCCCAC

     SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer
121  CTCGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCCACATC
     GAGCCCCCGCGAATACTGTATTATTAAACACTGCTCACGGTGAGGTGCCTACGGTGTAG

     IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuValVal
181  CATCTTGGGCATCGGCACTGTCCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGGTTGT
     GTAGAACCCGTAGCCGTGACAGGAACTGGTTCGTCTCTGACGCCCCCGCTCTGACCAACA

     LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal
241  GCTCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCCATCCCAACATCGAGGAGGT
     CGAGCGGTGGCGGTGGGGAGGCCCGAGGCAGTGACACGGGGTAGGGTTGTAGCTCCTCCA

     AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluValIle
301  TGCTCTGTCCACCACCGGAGAGATCCCTTTTTACGGCAAGGCTATCCCCCTCGAAGTAAT
     ACGAGACAGGTGGTGGCCTCTCTAGGGAAAAATGCCGTTCCGATAGGGGGAGCTTCATTA
                                                            --------

     LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla
361  CAAGGGGGGAGACATCTCATCTTCTGTCATTCAAAGAAGAAGTGCGACGAACTCGCCGC
     GTTCCCCCCTCTGTAGAGTAGAAGACAGTAAGTTTCTTCTTCACGCTGCTTGAGCGGCG
     ----------------------Overlap with 37b--------------------
     LysLeuValAlaLeuGlyIleAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal
421  AAAGCTGGTCGCATTGGGCATCAATGCCGTGGCCTACTACCGCGGTCTTGACGTGTCCGT
     TTTCGACCAGCGTAACCCGTAGTTACGGCACCGGATGATGGCGCCAGAACTGCACAGGCA
     ----------
     IleProThr
481  CATCCCGACCAG
     GTAGGGCTGGTC
```

Translation of DNA 40b FIG. 28

1 2 3

FIG. 29

HCV cDNA OF CLONE 40a

```
     GluPheGlyAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis
  1  GAATTCGGGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT
     CTTAAGCCCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

 61  TCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG
     AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

121  GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGTGATGTTGTCGTCGTG
     CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCACTACAACAGCAGCAC

181  GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT
     CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA

241  ACGTGTGTCACCCAGACAGTCGATTTCAGCCTTGACCCTACCTTCACTATTGAGACAATC
     TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAACTGGGATGGAAGTGATAACTCTGTTAG

301  ACGCTCCCCCAAGATGCTCCGAATTC
     TGCGAGGGGGTTCTACGAGGCTTAAG
```

FIG. 32

FIG. 33

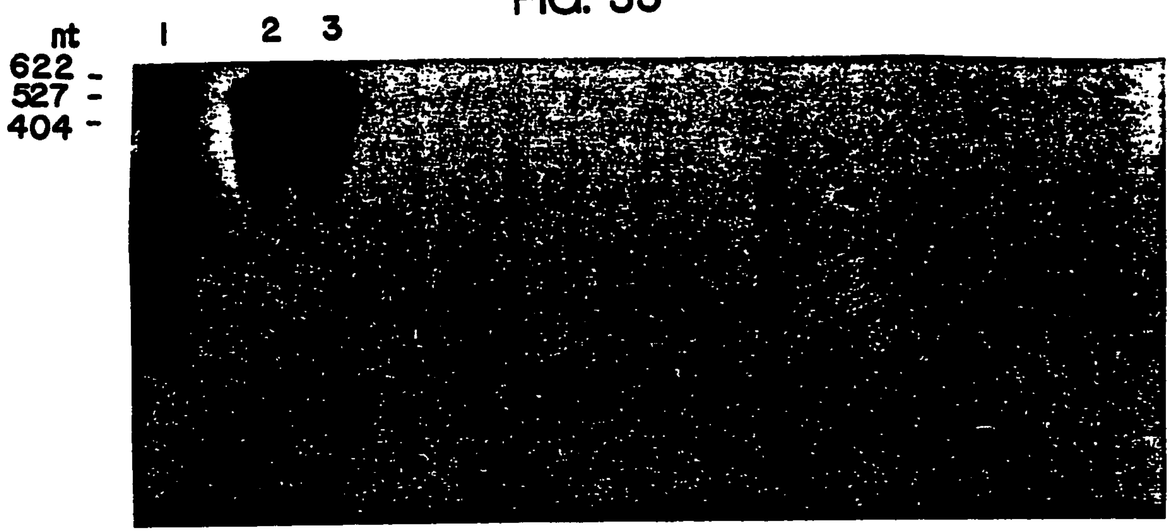


FIG. 34-1 Translation of DNA 35

```
    SerIleGluThrIleThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArg
  1 TCCATTGAGACAATCACGCTCCCCCAGGATGCTGTCTCCCGCACTCAACGTCGGGGCAGG
    AGGTAACTCTGTTAGTGCGAGGGGGTCCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCC

    ThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGlyGluArgProSerGly
 61 ACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGGGAGCGCCCCTCCGGC
    TGACCGTCCCCCTTCGGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCG

    MetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeu
121 ATGTTCGACTCGTCCGTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTC
    TACAAGCTGAGCAGGCAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAG

    ThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProVal
181 ACGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTG
    TGCGGGCGGCTCTGATGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCAC
```

FIG. 34-2

```
                                                                  ------
    CysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAla
241 TGCCAGGACCATCTTGAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCC
    ACGGTCCTGGTAGAACTTAAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGG

    ------------------------------------------------------------
    HisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGln
301 CACTTTCTATCCCAGACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAA
    GTGAAAGATAGGGTCTGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTT

    ----------------Overlap with 36-----------------------------
    AlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCys
361 GCCACCGTGTGCGCTAGGGCTCAAGCCCCTCCCCCATCGTGGACCAGATGTGGAAGTGT
    CGGTGGCACACGCGATCCCGAGTTCGGGGAGGGGGTAGCACCCTGGTCTACACCTTCACA

    ------------------------------------------------------------
    LeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAla
421 TTGATTCGCCTCAAGCCCACCCTCCATGGCCAACACCCCTGCTATACAGACTGGGCGCT
    AACTAAGCGGAGTTCGGGTGGGAGGTACCGGTTGTGGGGACGATATGTCTGACCCGCGA
```

NOT FURNISHED UPON FILING

NOT FURNISHED UPON FILING

PCR/HCV ENV Region

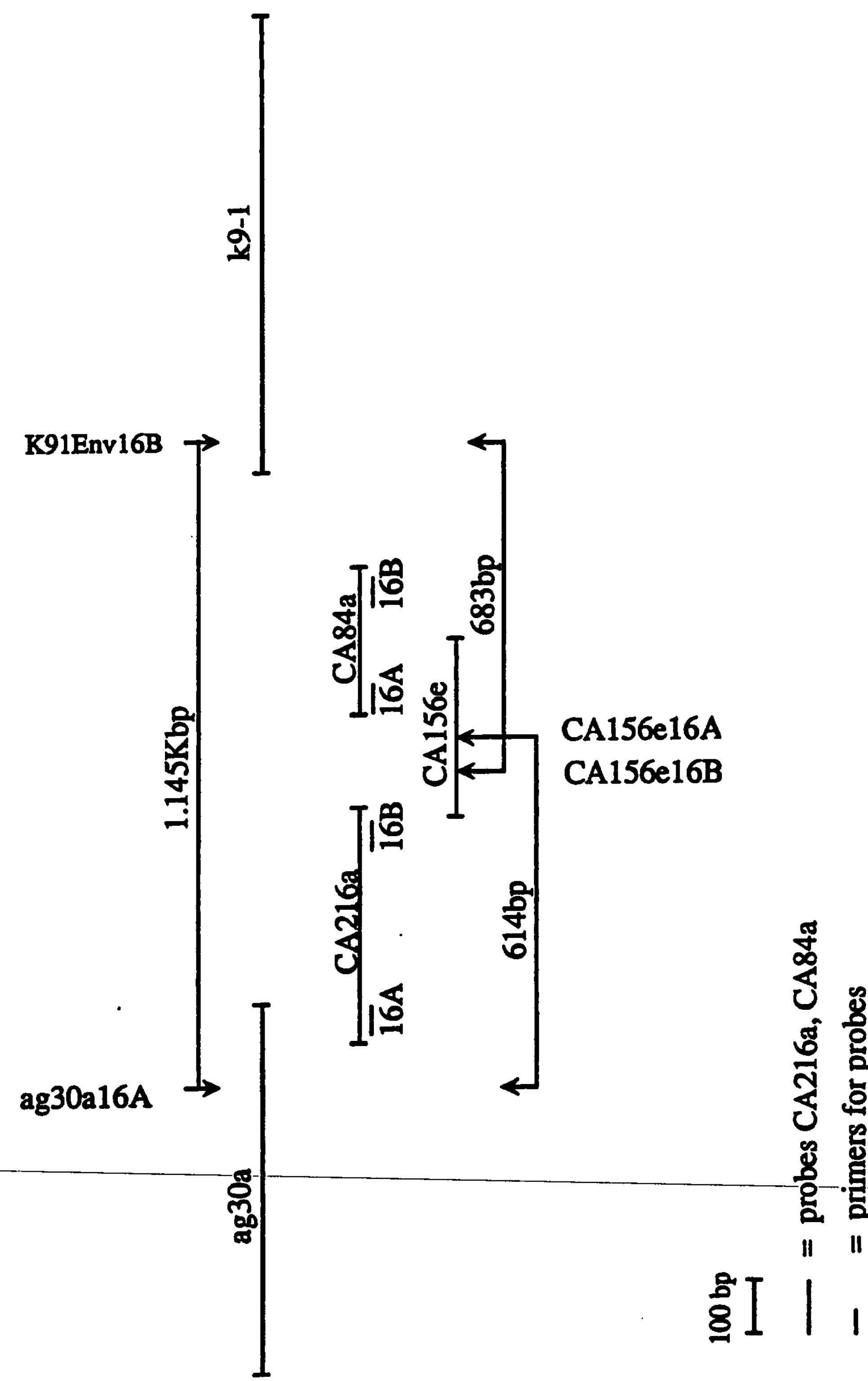



FIG. 37

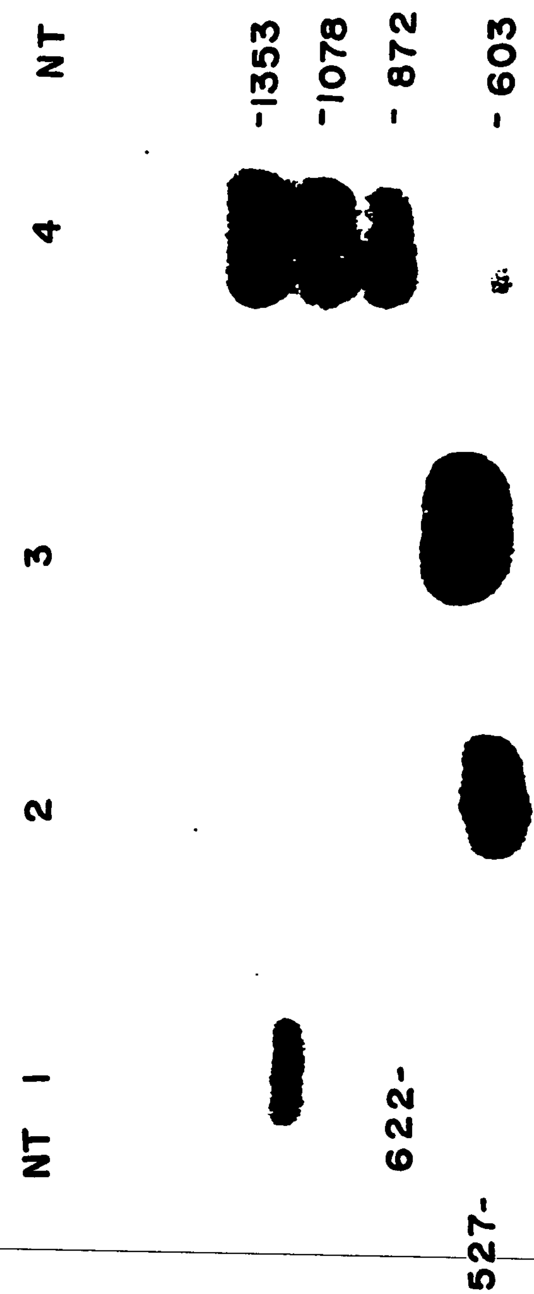



FIG. 38

FIG. 39-1

```
     1.  human 27     2.  HCV 1     3.  human 23

  1 CGGCTTCGCCGACCTCATGGGTACATtCCGCTCGTCGGCGCtCCTCTTGGgGGCGCTGCCAGGGCCCTGGC
    ************************** ************** ******** ********************
  1 CGGCTTCGCCGACCTCATGGGTACATACCGCTCGTCGGCGCCCCTCTTGGAGGCGCTGCCAGGGCCCTGGC
    ******************************************************  ***************
  1 CGGCTTCGCCGACCTCATGGGTACATACCGCTCGTCGGCGCCCCTCTTGGAGGCcgTGCCAGGGCCCTGGC

 73 GCATGGCGTCCGGGTTCTGGAAGACGGCGTGAACTATGCAACAGGGAACCTTCCTGGTTGCTCTTTCTCTAT
    ************************************************************************
 73 GCATGGCGTCCGGGTTCTGGAAGACGGCGTGAACTATGCAACAGGGAACCTTCCTGGTTGCTCTTTCTCTAT
    *** ************ ********************************************** ** *****
 73 GCAcGGCGTCCGGGTTtTGGAAGACGGCGTGAACTATGCAACAGGGAACCTTCCTGGTTGCTCcTTtTCTAT

145 CTTCCTTCTGGCtCTGCTCTCTTGCcTGACcGTGCCCGCaTCGGCCTACCAAGTaCGCAACTCCtCGGGcaT
    ************ ************ **** ******** ************** ********* ****  *
145 CTTCCTTCTGGCCCTGCTCTCTTGCtTGACtGTGCCCGCTTCGGCCTACCAAGTGCGCAACTCCACGGGGCT
    *************** ********* **** *********** ******************** ***** **
145 CTTCCTTCTGGCCCTaCTCTCTTGCcTGACcGTGCCCGCTTCaGCCTACCAAGTGCGCAACTCtACGGGGCT

217 TTACCAtGTCACCAATGATTGCCCTAAtTCGAGTATTGTGTACGAGaCGGCCGAcaCCATCCTaCACtCTCC
    ****** ******************** ****************** *******  ******* *** ****
217 TTACCAcGTCACCAATGATTGCCCTAACTCGAGTATTGTGTACGAGGCGGCCGATGCCATCCTGCACaCTCC
    ****** ************************************************************ ****
217 TTACCAtGTCACCAATGATTGCCCTAACTCGAGTATTGTGTACGAGGCGGCCGATGCCATCCTGCACgCTCC

289 GGGGTGtGTCCCTTGCGTTCGcGAGGGtAACGCCTCGAaaTGTTGGGTGcCGgTagCCCCcACaGTGGCCAC
    ****** ************** ***** **********  ********* ** *  **** ** ********
289 GGGGTGcGTCCCTTGCGTTCGtGAGGGcAACGCCTCGAGgTGTTGGGTGGCGaTGACCCCtACGGTGGCCAC
    ****** ************** ****  **** ****** ************ ******* ***********
289 GGGGTGtGTCCCTTGCGTTCGcGAGGatAACGtCTCGAGaTGTTGGGTGGCGgTGACCCCcACGGTGGCCAC
```

FIG. 39-2

```
361 CAGGGAcGGCAAcCTCCCCGCaACGCAGCTTCGACGTCACATCGATCTGCTTGTCGGGAGtGCCACCCTtTG
    ****** ***** ******** ************************************** ******** **
361 CAGGGAtGGCAAACTCCCCGCgACGCAGCTTCGACGTCACATCGATCTGCTTGTCGGGAGCGCCACCCTCTG
    ** *** ************ * ************************************************ *
361 CAaGGAcGGCAAACTCCCCaCaACGCAGCTTCGACGTCACATCGATCTGCTTGTCGGGAGCGCCACCCTCTG

433 cTCGGCCCTCTAtGTGGGGACtTGTGCGGGTCTGTCTTTCTTGTCGGtCAACTGTTCACtTTCTCcCCCAG
     *********** ********* ************************* *********** ***** ****
433 tTCGGCCCTCTACGTGGGGACCTGTGCGGGTCTGTCTTTCTTGTCGGcCAACTGTTCACCTTCTCTCCCAG
     ********************** ******** ************* ******** ***************
433 cTCGGCCCTCTACGTGGGGACCTtTGCGGGTCcaTCTTTCTTGTCGGtCAACTGTTtACCTTCTCTCCCAG
505 GCGCCACTGGACaACGCAAGaTTGCAAcTGCTCTATCTAcCCGGCCATATAACGGGaCACCGCATGGCATG
    ************ ******* ****** *********** ***************** ********* ****
505 GCGCCACTGGACGACGCAAGgTTGCAAtTGCTCTATCTATCCCGGCCATATAACGGGTCACCGCATGGCATG
    ****************** * ***** ** *****************************************
505 GCGCCACTGGACGACGCAgGacTGCAAcTGtTCTATCTATCCCGGCCATATAACGGGTCACCGCATGGCATG
577 GGATATGATGATGAACTGGTCCCCTACagCaGCGcTGGTAATGGCTCAGCTGCTCaGGATCCCgCAAGCCAT
    *************************** * *** ********************* ******* **** ***
577 GGATATGATGATGAACTGGTCCCCTACGaCGGCGTTGGTAATGGCTCAGCTGCTCCGGATCCCACAAGCCAT
    **************************** **** ****** * ************************** ***
577 GGATATGATGATGAACTGGTCCCCTACGgCGGCaTTGGTAgTaGCTCAGCTGCTCCGGATCCCACAAGCCAT
649 CTTGGACATGATCGCTGGTGCTCACTGGGGAGTCCTaGCGGGCATAGCGTATTTCTCCATGGTGGGAACTG
    *********************************** ************************************
649 CTTGGACATGATCGCTGGTGCTCACTGGGGAGTCCTGGCGGGCATAGCGTATTTCTCCATGGTGGGAACTG
    ********************************************** *************************
649 CTTGGACATGATCGCTGGTGCTCACTGGGGAGTCCTGGCGGGCATgGCGTATTTCTCCATGGTGGGAACTG
721 GGCGAAGGTCCTGGTgGTGCTGtTGCTgTTTGCCGGCGTCGAtGCGacAACCtAtacCACCGGGGGgAaTGC
    *************** ****** **** ************** *** **** * ********* * ***
721 GGCGAAGGTCCTGGTAGTGCTGCTGCTATTTGCCGGCGTCGACGCGGAAACCCACgtCACCGGGGGAAGTGC
    ********** ************* ************* ** ************* * ***********
721 GGCGAAGGTCCTGGTAGTGCTGCTtCTATTTGCCGGCGTCGACGCGGAAACCCACcgtACCGGGGGAAGTGC
```

SUBSTITUTE SHEET

```
793  tGcCaggACcacGcagGcgcTcaccAGttTttTCagcCCAGGCGCCAAGCAGgAtaTCCAGCTGATCAACAC
      * *   **   *   *   *    **  *  **   *************** *  ****************
793  CGgCCaCACtgtGtCTGGAtTTGtTAGcCTCcTCgCACCAGGCGCCAAGCAGAACgTCCAGCTGATCAACAC
     ** ** **    * ***** *** *** *** ** ********** * ******* ****************
793  CGcCCgCAgcacGgCTGGAgTTGcTAGtCTCtTCaCACCAGGCGCtAgGCAGAACaTCCAGCTGATCAACAC

865  CAACGGCAGTTGGCACaTCAATcGCACGGCCtTGAACTGtAATGcgAGCCTCgACACtGGCTGGgTaGCgGG
     **************** ***** ******** ******* ****  ****** **** ****** * ** **
865  CAACGGCAGTTGGCACcTCAATAGCACGGCCcTGAACTGCAATGAtAGCCTCAACACCGGCTGGTTgGCaGG
     **************** ******* ****** ************* ***** * ************ ** **
865  CAACGGCAGTTGGCACaTCAATAGtACGGCCtTGAACTGCAATGAcAGCCTtAcCACCGGCTGGTTaGCgGG

937  GCTcTTCTATtACCACAAaTTCAACTCTTCAGGCTGcCCcGAGAGGaTgGCCAGCTGtaGgCCCCTTgCCGA
     *** ****** ******* ***************** ** ****** * ********  * ****** ****
937  GCTTTTCTATCACCACAAgTTCAACTCTTCAGGCTGTCCtGAGAGGcTaGCCAGCTGCCGACCCCTTACCGA
     *************** ** ******************** ****** * ***************** *****
937  GCTTTTCTATCACCAtAAaTTCAACTCTTCAGGCTGTCCcGAGAGGtTgGCCAGCTGCCGACCCCTcACCGA

1009 TTTcGACCAGG
     *** *******
1009 TTTTGACCAGG
     ***** *****
1009 TTTTGcCCAGG
```

FIG. 39-3

FIG. 40

```
  1 GFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLALLSCLTVPASAYQVRNSsGi
    ********************************************************************* * 
  1 GFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLALLSCLTVPASAYQVRNSTGL
    ****************** *****************************************************
  1 GFADLMGYIPLVGAPLGGrARALAHGVRVLEDGVNYATGNLPGCSFSIFLLALLSCLTVPASAYQVRNSTGL
 73 YHVTNDCPNSSIVYEtADtILHsPGCVPCVREGNASkCWVpvaPTVATRDGnLPATQLRRHIDLLVGSATLC
    *************** ** *** ************* ***   ******** ********************
 73 YHVTNDCPNSSIVYEAADAILHtPGCVPCVREGNASRCWVAmTPTVATRDGKLPATQLRRHIDLLVGSATLC
    ********************** ********* * ****** ****** ***** *****************
 73 YHVTNDCPNSSIVYEAADAILHaPGCVPCVREdNvSRCWVAvTPTVATkDGKLPtTQLRRHIDLLVGSATLC
145 SALYVGDLCGSVFLVGQLFTFSPRRHWTTQdCNCSIYPGHITGHRMAWDMMMNWSPTaALVMAQLLRIPQAI
    ****************************** ************************** **************
145 SALYVGDLCGSVFLVGQLFTFSPRRHWTTQgCNCSIYPGHITGHRMAWDMMMNWSPTtALVMAQLLRIPQAI
    *********** ****************** ************************** *** **********
145 SALYVGDLCGSiFLVGQLFTFSPRRHWTTQdCNCSIYPGHITGHRMAWDMMMNWSPTaALVvAQLLRIPQAI
217 LDMIAGAHWGVLAGIAYFSMVGNWAKVLVVLLLFAGVDAtTytTGGnAarTtqaltSffsPGAKQdiQLINT
    *************************************** *  *** *  *     *   *****  *****
217 LDMIAGAHWGVLAGIAYFSMVGNWAKVLVVLLLFAGVDAETHvTGGSAghTvsGfvSLlaPGAKQNvQLINT
    ************** *************************** *****     *  **  *** ** *****
217 LDMIAGAHWGVLAGmAYFSMVGNWAKVLVVLLLFAGVDAETHrTGGSAarstaGvaSLftPGArQNiQLINT
289 NGSWHiNrTALNCNaSLdTGWvAGLFYyHKFNSSGCPERmASCRPLaDFDQ
    ***** * ****** ** *** ***** *********** ****** ****
289 NGSWHlNSTALNCNDSLnTGWLAGLFYHHKFNSSGCPERLASCRPLTDFDQ
    ***** *********** ******************************* *
289 NGSWHiNSTALNCNDSLtTGWLAGLFYHHKFNSSGCPERLASCRPLTDFaQ
```

1. human 27
2. HCV 1
3. human 23

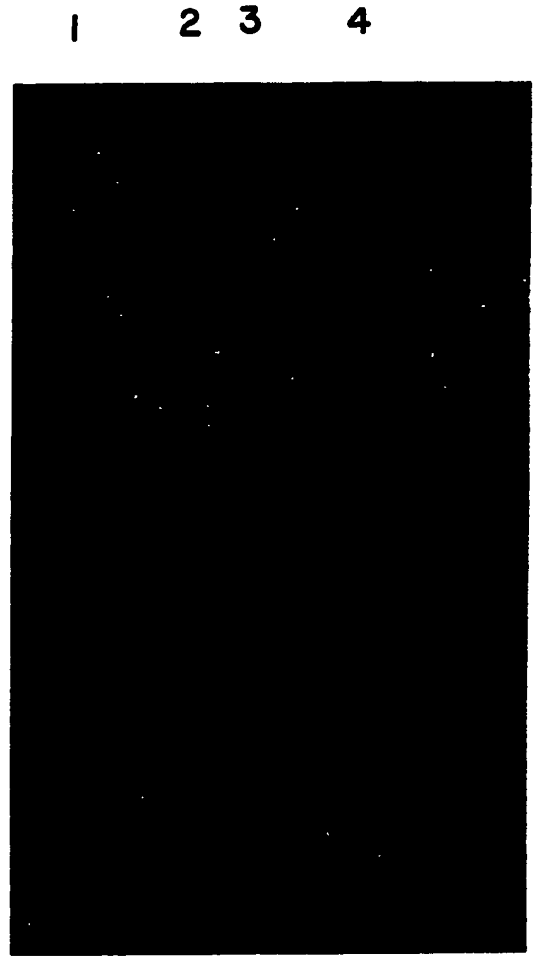



FIG. 41

1 2 3 4

A

B

C

FIG. 43

A B C

1

2

3

4

FIG. 44A

A B C

1

2

3

4

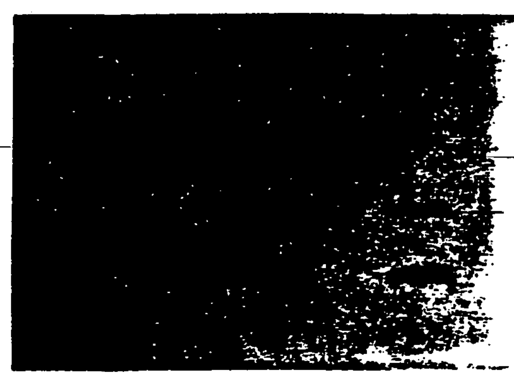


FIG. 44B

FIG. 46-1

Human 23

```
     GlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeuGlyGlyArgAla
  1  GGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCGGCGCCCCTCTTGGAGGCCGTGCC

     ArgAlaLeuAlaHisGlyValArgValLeuGluAspGlyValAsnTyrAlaThrGlyAsn
 61  AGGGCCCTGGCGCACGGCGTCCGGGTTTTGGAAGACGGCGTGAACTATGCAACAGGGAAC
              CG   A
     LeuProGlyCysSerPheSerIlePheLeuLeuAlaLeuLeuSerCysLeuThrValPro
121  CTTCCTGGTTGCTCCTTTTCTATCTTCCTTCTGGCCCTACTCTCTTGCCTGACCGTGCCC
                        GA                    T
     AlaSerAlaTyrGlnValArgAsnSerThrGlyLeuTyrHisValThrAsnAspCysPro
181  GCTTCAGCCTACCAAGTGCGCAACTCTACGGGGCTTTACCATGTCACCAATGATTGCCCT
     AsnSerSerIleValTyrGluAlaAlaAspAlaIleLeuHisAlaProGlyCysValPro
241  AACTCGAGTATTGTGTACGAGGCGGCCGATGCCATCCTGCACGCTCCGGGGTGTGTCCCT
                           T                            C
     CysValArgGluAspAsnValSerArgCysTrpValAlaValThrProThrValAlaThr
301  TGCGTTCGCGAGGATAACGTCTCGAGATGTTGGGTGGCGGTGACCCCCACGGTGGCCACC
             G                                                 T
     LysAspGlyLysLeuProThrThrGlnLeuArgArgHisIleAspLeuLeuValGlySer
361  AAGGACGGCAAACTCCCCACAACGCAGCTTCGACGTCACATCGATCTGCTTGTCGGGAGC
                                                       C       A
     AlaThrLeuCysSerAlaLeuTyrValGlyAspLeuCysGlySerIlePheLeuValGly
421  GCCACCCTCTGCTCGGCCCTCTACGTGGGGGACCTTTGCGGGTCCATCTTTCTTGTCGGT
             T

     GlnLeuPheThrPheSerProArgArgHisTrpThrThrGlnAspCysAsnCysSerIle
481  CAACTGTTTACCTTCTCTCCCAGGCGCCACTGGACGACGCAGGACTGCAACTGTTCTATC
                                             C
```

FIG. 46-2

```
     TyrProGlyHisIleThrGlyHisArgMetAlaTrpAspMetMetMetAsnTrpSerPro
541  TATCCCGGCCATATAACGGGTCACCGCATGGCATGGGATATGATGATGAACTGGTCCCCT
                                                      G

     ThrAlaAlaLeuValValAlaGlnLeuLeuArgIleProGlnAlaIleLeuAspMetIle
601  ACGGCGGCATTGGTAGTAGCTCAGCTGCTCCGGATCCCACAAGCCATCTTGGACATGATC
                   G   AG

     AlaGlyAlaHisTrpGlyValLeuAlaGlyMetAlaTyrPheSerMetValGlyAsnTrp
661  GCTGGTGCTCACTGGGGAGTCCTGGCGGGCATGGCGTATTTCTCCATGGTGGGAACTGG
                                            G

     AlaLysValLeuValValLeuLeuLeuPheAlaGlyValAspAlaGluThrHisArgThr
721  GCGAAGGTCCTGGTAGTGCTGCTTCTATTTGCCGGCGTCGACGCGGAAACCCACCGTACC
                                                G

     GlyGlySerAlaAlaArgSerThrAlaGlyValAlaSerLeuPheThrProGlyAlaArg
781  GGGGGAAGTGCCGCCCGCAGCACGGCTGGAGTTGCTAGTCTCTTCACACCAGGCGCTAGG
     C          T   A

     GlnAsnIleGlnLeuIleAsnThrAsnGlySerTrpHisIleAsnSerThrAlaLeuAsn
841  CAGAACATCCAGCTGATCAACACCAACGGCAGTTGGCACATCAATAGTACGGCCTTGAAC
                                                    AT

     CysAsnAspSerLeuThrThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn
901  TGCAATGACAGCCTTACCACCGGCTGGTTAGCGGGGCTTTTCTATCACCATAAATTCAAC
       A                             A

     SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAlaGln
961  TCTTCAGGCTGTCCCGAGAGGTTGGCCAGCTGCCGACCCCTCACCGATTTTGCCCAGG
      G  A                                    G
```

FIG. 47-1

Human 27

```
       GlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeuGlyGlyAlaAla
    1  GGCTTCGCCGACCTCATGGGTACATTCCGCTCGTCGGCGCTCCTCTTGGGGGCGCTGCC

       ArgAlaLeuAlaHisGlyValArgValLeuGluAspGlyValAsnTyrAlaThrGlyAsn
   61  AGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAAGACGGCGTGAACTATGCAACAGGGAAC

       LeuProGlyCysSerPheSerIlePheLeuLeuAlaLeuLeuSerCysLeuThrValPro
  121  CTTCCTGGTTGCTCTTTCTCTATCTTCCTTCTGGCTCTGCTCTCTTGCCTGACCGTGCCC

       AlaSerAlaTyrGlnValArgAsnSerSerGlyIleTyrHisValThrAsnAspCysPro
  181  GCATCGGCCTACCAAGTACGCAACTCCTCGGGCATTTACCATGTCACCAATGATTGCCCT

       AsnSerSerIleValTyrGluThrAlaAspThrIleLeuHisSerProGlyCysValPro
  241  AATTCGAGTATTGTGTACGAGACGGCCGACACCATCCTACACTCTCCGGGGTGTGTCCCT
          C           C
       CysValArgGluGlyAsnAlaSerLysCysTrpValProValAlaProThrValAlaThr
  301  TGCGTTCGCGAGGGTAACGCCTCGAAATGTTGGGTGCCGGTAGCCCCCACAGTGGCCACC
                 G
       ArgAspGlyAsnLeuProAlaThrGlnLeuArgArgHisIleAspLeuLeuValGlySer
  361  AGGGACGGCAACCTCCCCGCAACGCAGCTTCGACGTCACATCGATCTGCTTGTCGGGAGT
                                       G          G

       AlaThrLeuCysSerAlaLeuTyrValGlyAspLeuCysGlySerValPheLeuValGly
  421  GCCACCCTTTGCTCGGCCCTCTATGTGGGGGACTTGTGCGGGTCTGTCTTTCTTGTCGGT
               C                                         C
       GlnLeuPheThrPheSerProArgArgHisTrpThrThrGlnAspCysAsnCysSerIle
  481  CAACTGTTCACTTTCTCCCCCAGGCGCCACTGGACAACGCAAGATTGCAACTGCTCTATC
                        A
```

FIG. 47-2

```
     TyrProGlyHisIleThrGlyHisArgMetAlaTrpAspMetMetMetAsnTrpSerPro
541  TACCCCGGCCATATAACGGGACACCGCATGGCATGGGATATGATGATGAACTGGTCCCCT

     ThrAlaAlaLeuValMetAlaGlnLeuLeuArgIleProGlnAlaIleLeuAspMetIle
601  ACAGCAGCGCTGGTAATGGCTCAGCTGCTCAGGATCCCGCAAGCCATCTTGGACATGATC
          G

     AlaGlyAlaHisTrpGlyValLeuAlaGlyIleAlaTyrPheSerMetValGlyAsnTrp
661  GCTGGTGCTCACTGGGGAGTCCTAGCGGGCATAGCGTATTTCTCCATGGTGGGAACTGG

     AlaLysValLeuValValLeuLeuLeuPheAlaGlyValAspAlaThrThrTyrThrThr
721  GCGAAGGTCCTGGTGGTGCTGTTGCTGTTTGCCGGCGTCGATGCGACAACCTATACCACC

     GlyGlyAsnAlaAlaArgThrThrGlnAlaLeuThrSerPhePheSerProGlyAlaLys
781  GGGGGGAATGCTGCCAGGACCACGCAGGCGCTCACCAGTTTTTTCAGCCCAGGCGCCAAG

     GlnAspIleGlnLeuIleAsnThrAsnGlySerTrpHisIleAsnArgThrAlaLeuAsn
841  CAGGATATCCAGCTGATCAACACCAACGGCAGTTGGCACATCAATCGCACGGCCTTGAAC
      G                         T

     CysAsnAlaSerLeuAspThrGlyTrpValAlaGlyLeuPheTyrTyrHisLysPheAsn
901  TGTAATGCGAGCCTCGACACTGGCTGGGTAGCGGGGCTCTTCTATTACCACAAATTCAAC
                 T                G

     SerSerGlyCysProGluArgMetAlaSerCysArgProLeuAlaAspPheAspGln
961  TCTTCAGGCTGCCCCGAGAGGATGGCCAGCTGTAGGCCCCTTGCCGATTTCGACCAGG
                                                    C
```

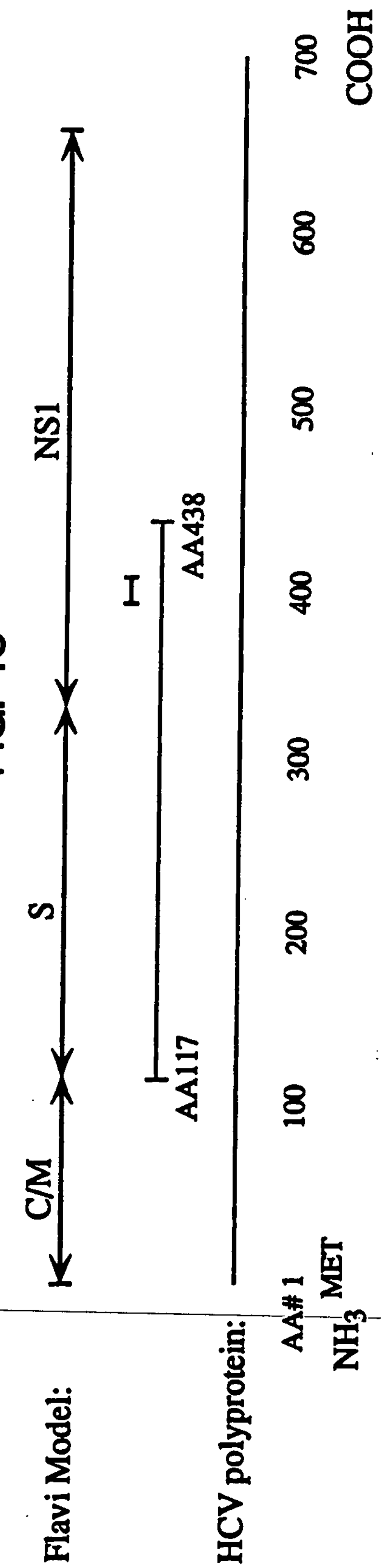

Flavi Model:
C/M
S
NS1
HCV polyprotein:
AA117
AA438
I
AA# 1
$NH_3$ MET
100
200
300
400
500
600
700
COOH

FIG. 48

CLUSTERED PAIR-WISE 'REGION' ALIGNMENT
in 'identity (no translation)' alphabet of:

FIG. 49-1

1. ssThorn#8.r (1-587)
2. ssEC1#2.r (1-587)
3. ssHCT18#7.r (1-587)
4. env1.hcv (1-1657)

```
                                                                               GA
                                                                               ||
                                                                               GA
                                                                               ||
                                                                               GA
                                                                               |
289 gggtgggcgggatggctcctgtctccccgtggctctcggcctagctggggccccacagacccccggcgtaGg

  3 ATTCGCAATTTGGGTAAGGTCATCGATACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTC
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  3 ATTCGCAATTTGGGTAAGGTCATCGATACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTC
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
  3 ATTCGCAATTTGGGTAAGGTCATCGATACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTAtATACCGCTC
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
361 tcgCGCAATTTGGGTAAGGTCATCGATACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTAcATACCGCTC

 75 GTCGGCGCCCCTCTTGGgGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAAGACGGCGTGAAC
    ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
 75 GTCGGCGCCCCTCTTGGAGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAAGACGGCGTGAAC
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 75 GTCGGCGCCCCTCTTGGAGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAAGACGGCGTGAAC
    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
433 GTCGGCGCCCCTCTTGGAGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAAGACGGCGTGAAC
```

FIG. 49-2

```
147 TATGCAACAGGGAACCTTCCTGGTTGCTCTTTCTcTcTCTTCCTTCTGGCCcTGCTCTCTTGtcTGACcGTG
147 TATGCAACAGGGAACCTTCCTGGTTGCTCTTTCTtTATCTTCCTTCTGGCCtTGCTCTCTTGCtTGACTGTG
147 TATGC   CAGGGAACCTTCCTGGTTGCTCTTTCTCTATCTTCCTTCTGGCCCTGCTCTCTTGCcTGACTGTG
505 TATGCAACAGGGAACCTTCCTGGTTGCTCTTTCTCTATCTTCCTTCTGGCCCTGCTCTCTTGCtTGACTGTG
219 CCCGCTTCAGCCTACCAAGTGCGCAACTCCaCGGGGCTTTACCATGTCACCAAcGATTGCCCcAACTCGAGt
219 CCCGCTTCAGCCTACCAAGTGCGCAACTCCtCGGGGCTTTACCATGTCACCAATGATTGCCCtAACTCGAGc
219 CCCGCTTCAGCCcACCAAGTGCGCAACTCCACGGGGCTTTACCATGTCACCAATGATTGCCCcAACTCGAGT
577 CCCGCTTCgGCCtACCAAGTGCGCAACTCCACGGGGCTTTACCAcGTCACCAATGATTGCCCtAACTCGAGT
291 ATTGTGTACGAGGCGGCCGATGCtATCCTGCACgCTCCGGGGTGTGTCCCTTGCGTTCgCGAGGGtAACGcC
291 ATTGTGTACGAGGCGGCCGATGCCATCCTGCACACTCCGGGGTGTGTCCCTTGCGTTCACGAGGGCAACGTC
291 ATTGTaTACGAaGCGGCCGAcGCCATCCTGCACACTCCGGGGTGTGTCCCTTGCGTTCACGAGGGCAACGTC
649 ATTGTgTACGAgGCGGCCGAtGCCATCCTGCACACTCCGGGGTGcGTCCCTTGCGTTCgtGAGGGCAACGcC
363 TCGAGGTGTTGGGTGGCGATGACCCCCACGGTGGCCgCCAGGGaCGGCAgACTCCCCACAACGCAGCTgCGA
363 TCGAGGTGTTGGGTGGCGATGACCCCCACGGTGGCCACCAGGGgCGGCAAACTCCCCACAACGCAGCTTCGA
363 TCGAGGTGTTGGGTGGCGgTGACCCCCACGGTGGCCACCAGGGATGGCAAACTCCCCACAACGCAGCTTCGA
721 TCGAGGTGTTGGGTGGCGaTGACCCCtACGGTGGCCACCAGGGATGGCAAACTCCCCgCgACGCAGCTTCGA
```

```
435 CGTCACATCGATCTGCTTGTCGGGAGCGCcACCCTCTGCTCGGCCCTCTACGTGGGGACCTGTGCGGGTCc
435 CGTCACATCGATCTGCTTGTCGGGAGCGCtACCCTCTGCTCGGCCCTCTACGTGGGGACCTGTGCGGGTCT
435 CGTCACATCGATCTGCTTGTCGGGAGCGCCACCCTCTGCTCGGCCCTCTAtGTGGGGACtTGTGCGGGTCT
793 CGTCACATCGATCTGCTTGTCGGGAGCGCCACCCTCTGtTCGGCCCTCTAcGTGGGGACcTaTGCGGGTCT

507 aTCTTtCTTGTCGGTCAACTGTTcACCTTCTCTCCCAGGCGCCACTGGACGACGCAAGGTTGCAATTGCTCT
507 GTCTTcCTTGTCGGTCAACTGTTTACCTTCTCTCCCAGGCGCCACTGGACGACGCAAGGTTGCAATTGCTCT
507 GTCTTTCTTGTCGGCCAACTGTTTACCTTCTCTCCCAGGCGCCACTGGACGACGCAAGGTTGCAATTGCTCT
865 GTCTTTCTTGTCGGCCAACTGTTcACCTTCTCTCCCAGGCGCCACTGGACGACGCAAGGTTGCAATTGCTCT

579 ATCGAATTC
579 ATCGAATTC
579 ATCGAATTC
937 ATCtAtccC
```

FIG. 49-3

```
                         10        20        30        40
EC10          GAATTCGGACGACGCAAGGTTGCAATTGCTCTATCTATCCCGGCCATAT
                   X::::::::::::::::::::::::::::::::::::::::::
HCV1   CTCTCCCAGGCGCCACTGGACGACGCAAGGTTGCAATTGCTCTATCTATCCCGGCCATAT
              550       560       570       580       590       600

       50        60        70        80 A      90        100
       AACAGGTCACCGCATGGCATGGGATATGATGATGAACTGGTCCCCTACGACGGCGTTAGT
       ::: :::::::::::::::::::::::::::::::::::::::::::::::::::: ::
       AACGGGTCACCGCATGGCATGGGATATGATGATGAACTGGTCCCCTACGACGGCGTTGGT
              610       620       630       640       650       660

      110       120       130       140       150       160
       GGTAGCTCAGCTGCTCCGGATCCCACAAGCCATCTTGGACATGATCGCTGGTGCTCACTG
         : ::::::::::::::::::::::::::::::::::::::::::::::::::::::::
       AATGGCTCAGCTGCTCCGGATCCCACAAGCCATCTTGGACATGATCGCTGGTGCTCACTG
              670       680       690       700       710       720

      170       180       190       200       210       220
       GGGAGTCCTGGCGGGCATAGCGTATTTCTCCATGGTGGGGAACTGGGCGAAGGTCTTGGC
       :::::::::::::::::::::::::::::::::::::::::::::::::::::: :::
       GGGAGTCCTGGCGGGCATAGCGTATTTCTCCATGGTGGGGAACTGGGCGAAGGTCCTGGT
              730       740       750       760       770       780

      230       240       250       260       270       280
       AGTGCTGCTGCTATTTGCCGGCGTCGACGCGGAAACCCACGTCACTGGGGGGATCGCCGC
       :::::::::::::::::::::::::::::::::::::::::::: ::::: :  ::::
       AGTGCTGCTGCTATTTGCCGGCGTCGACGCGGAAACCCACGTCACCGGGGGAAGTGCCGG
              790       800       810       820       830       840

      290       300       310       320       330       340
       CAAAACTACGGCTAGCCTTACTGGTCTCTTCAATTTAGGTGCCAAGCAGAACATCCAGCT
       : : :::  : :: :   ::  :  : ::: ::     ::: ::::::::::: :::::::
       CCACACTGTGTCTGGATTTGTTAGCCTCCTCGCACCAGGCGCCAAGCAGAACGTCCAGCT
              850       860       870       880       890       900

      350       360       370       380       390       400
       GATCAACACCAACGGCAGTTGGCACATCAACAGGACGGCCTTGAACTGCAATGATAGCCT
       :::::::::::::::::::::::::: :::: :: :::::: ::::::::::::::::::
       GATCAACACCAACGGCAGTTGGCACCTCAATAGCACGGCCCTGAACTGCAATGATAGCCT
              910       920       930       940       950       960

      410       420
       CAACACCGGCTGGAATTC
       ::::::::::::X
       CAACACCGGCTGGTTGGCAGGGCTTTTCTATCACCACAAGTTCAACTCTTCAGGCTGTCC
              970       980       990      1000      1010      1020
```

FIG. 50-1

NOT FURNISHED UPON FILING

FIG. 51

AA #117-308 (putative envelope region)

1) HCT #18 (USA) 3 clones sequenced
2) JH23 (USA) ?
3) JH 27 (USA) ?
4) PBL-Th (USA) 2 clones sequenced
5) EC1 (Italy) 3 clones sequenced
6) HCV-1 (chimpanzee) multiple

```
   C/M<-T->S
1)                     (P)
2)
3)
4)
5)
6)RNLGKVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYATGNL

1)                      H
2)
3)                            S                 T   T
4)       L
5)      (F)
6)PGCSFSIFLLALLSCLTVPASAYQVRNSTGLYHVTNDCPNSSIVYEAADAILH

           Y
1)        (H)   V       V           T
2)A          D  V       V     K     T
3)S                   PVA       N
4)A                         A   R   T
5)         H    V                   T
6)TPGCVPCVREGNASRCWVAMTPTVATRDGKLPATQLRRHIDLLVGSATLCS

1)
2)           I                  D
3)                              D
4)
5)           I
6)ALYVGDLCGSVFLVGQLFTFSPRRHWTTQGCNCSI
```

SUMMARY: "S" AA117-308 (93%)

HCT#18, PBL-Th, EC1(Italy) have 97% homology with HCV-1

JH23 and JH 27 have 96% and 95% hom logy with HCV-1,respectively

<u>AA#300-438 ( C-terminal region of the putative envelope region and amino ~1/3 of NSI)</u>

| | |
|---|---|
| 1) JH23 | ? |
| 2) JH27 | ? |
| 3) Japanese Isolate (T. Miyamura) | ? |
| 4) EC10 (Italy) | 2 clones sequenced (one nt difference, which did not result in an amino acid change) |
| 5) HCV-1 (chimpanzee) | multiple |

```
                                 S<-->NS I
1)   D                           A   V
2)   D                           A
3)                                   VS        VM V
4)
5)TTQGCNCSIYPGHITGHRMAWDMMMNWSPTTALVMAQLLRIPQAILDMIAGA

1)        M                          R      ARSTA VA
2)                                T YT    N AR TQALT F
3)        L Y           I M       GH R     VQ  VT TLT
4)                      A                 I AK TASLTA
5)HWGVLAGIAYFSMVGNWAKVLVVLLLFAGVDAETHVTGGSAGHTVSGFVSL

1)FS   R  I          I           T   V
2)FT       DI         I R      A  D
3)FR   S KI  V       I R         Q  F
4)FNL     I          I R             N
5)LAPGAKQNVQLINTNGSWHLNSTALNCNDSLNTGWL
```

SUMMARY: NS 1 AA 330-660

| "Isolate" | %Homology (AA330-438) | %Homology (AA383-405) |
|---|---|---|
| JH23 | 83 | 57 |
| JH27 | 80 | 39 |
| Japanese | 73 | 48 |
| EC10 (Italy) | 84 | 48 |

FIG. 52

FIG. 53

| Name | Common Sequence | Variable Sequence |
|---|---|---|
| 5'-3-1 | AAGCTTGATCGAATTC | CGATCTTGC |
| -2 | | CGATCCTGC |
| -3 | | CGATCATGC |
| -4 | | CGATCGTGC |
| -5 | | CGAAGTTGC |
| -6 | | CGAAGCTGC |
| -7 | | AGATCTTGC |
| -8 | | AGATCCTGC |
| -9 | | AGATCATGC |
| -10 | | AGATCGTGC |
| -11 | | AGAAGTTGC |
| -12 | | AGAAGCTGC |
| -13 | | CGATCTTGT |
| -14 | | CGATCCTGT |
| -15 | | CGATCATGT |
| -16 | | CGATCGTGT |
| -17 | | CGAAGTTGT |
| -18 | | CGAAGCTGT |
| -19 | | AGATCTTGT |
| -20 | | AGATCCTGT |
| -21 | | AGATCATGT |
| -22 | | AGATCGTGT |
| -23 | | AGAAGTTGT |
| -24 | | AGAAGCTGT |
| -25 | | CGCTCTTGC |
| -26 | | CGCTCCTGC |
| -27 | | CGCTCATGC |
| -28 | | CGCTCGTGC |
| -29 | | CGCAGTTGC |
| -30 | | CGCAGCTGC |
| -31 | | CGCTCTTGT |
| -32 | | CGCTCCTGT |
| -33 | | CGCTCATGT |
| -34 | | CGCTCGTGT |
| -35 | | CGCAGTTGT |
| -36 | | CGCAGCTGT |

# INTERNATIONAL SEARCH REPORT

International Application No PCT/US90/02853

**I. CLASSIFICATI N OF SUBJECT MATTER** (if several classification symbols apply, indicate all) [3]

According to International Patent Classification (IPC) or to both National Classification and IPC
IPC (5): C12Q 1/20
U.S. CL.: 435/5

**II. FIELDS SEARCHED**

Minimum Documentation Searched [4]

| Classification System | Classification Symbols |
|---|---|
| U.S. | 435/5 |

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [5]

Databases: USPTO Automated Patent System (File U.S. Pat. 1925-90)
Genebank UEMBL

**III. DOCUMENTS CONSIDERED TO BE RELEVANT** [14]

| Category * | Citation of Document, [16] with indication, where appropriate, of the relevant passages [17] | Relevant to Claim No. [18] |
|---|---|---|
| X | US, A, 4,683,195 (MULLIS ET AL) 28 July 1987 See the entire document. | 16, 17, 19 |
| X | US, A, 4,683,202 (MULLIS ET AL) 28 July 1987 See the entire document. | 16, 17, 19 |

* Special categories of cited documents: [15]
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

**IV. CERTIFICATION**

| Date of the Actual Completion of the International Search [2] | Date of Mailing of this International Search Report [2] |
|---|---|
| 22 AUGUST 1990 | 28 SEP 1990 |
| **International Searching Authority [1]** | **Signature of Authorized Officer [20]** |
| ISA/US | BRADLEY L. SISSON |

Form PCT/ISA/210 (second sheet) (May 1986)